The Continual Intercomparison of Radiation Codes: Results from Phase I

Lazaros Oreopoulos,[1] Eli Mlawer,[2] Jennifer Delamere,[2,3] Timothy Shippert,[4] Jason Cole,[5] Boris Fomin,[6] Michael Iacono,[2] Zhonghai Jin,[7,8] Jiangnan Li,[5] James Manners,[9] Petri Räisänen,[10] Fred Rose,[7,8] Yuanchong Zhang,[11] Michael J. Wilson,[1,12,13] and William B. Rossow[14]

[1]   We present results from Phase I of the Continual Intercomparison of Radiation Codes (CIRC), intended as an evolving and regularly updated reference source for evaluation of radiative transfer (RT) codes used in global climate models and other atmospheric applications. CIRC differs from previous intercomparisons in that it relies on an observationally validated catalog of cases. The seven CIRC Phase I baseline cases, five cloud free and two with overcast liquid clouds, are built around observations by the Atmospheric Radiation Measurements program that satisfy the goals of Phase I, namely, to examine RT model performance in realistic, yet not overly complex, atmospheric conditions. Besides the seven baseline cases, additional idealized "subcases" are also employed to facilitate interpretation of model errors. In addition to quantifying individual model performance with respect to reference line-by-line calculations, we also highlight RT code behavior for conditions of doubled CO_2, issues arising from spectral specification of surface albedo, and the impact of cloud scattering in the thermal infrared. Our analysis suggests that improvements in the calculation of diffuse shortwave flux, shortwave absorption, and shortwave CO_2 forcing as well as in the treatment of spectral surface albedo should be considered for many RT codes. On the other hand, longwave calculations are generally in agreement with the reference results. By expanding the range of conditions under which participating codes are tested, future CIRC phases will hopefully allow even more rigorous examination of RT codes.

Citation: Oreopoulos, L., et al. (2012), The Continual Intercomparison of Radiation Codes: Results from Phase I, *J. Geophys. Res.*, *117*, D06118, doi:10.1029/2011JD016821.

[1]Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

[2]Atmospheric and Environmental Research Inc., Lexington, Massachusetts, USA.

[3]Now at Tech-X, Boulder, Colorado, USA.

[4]Pacific Northwest National Laboratory, Richland, Washington, USA.

[5]Canadian Centre for Modeling and Analysis, Environment Canada, Victoria, British Columbia, Canada.

[6]Central Aerological Observatory, Dolgoprudny, Russia.

[7]Science Systems and Applications Inc., Hampton, Virginia, USA.

[8]Science Directorate, NASA Langley Research Center, Hampton, Virginia, USA.

[9]Met Office, Exeter, UK.

[10]Finnish Meteorological Institute, Helsinki, Finland.

[11]Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, USA.

[12]Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, Maryland, USA.

[13]Now at I.M. Systems Group, Rockville, Maryland, USA.

[14]Cooperative Remote Sensing Science and Technology Center, City College of New York, New York, New York, USA.

1.   Introduction

[2] While careful comparisons with observations have instilled us with a high degree of confidence in the gaseous radiative forcing calculations of high-resolution spectral radiation codes [*Mlawer et al.*, 2000; *Turner et al.*, 2004], their steep computational cost makes them presently unaffordable in global climate models (GCMs). Calculations from such codes have instead been used as the basis for designing approximate but much faster codes to perform efficient gaseous radiative transfer (RT) in GCMs. When clouds, aerosols and reflective/emitting surfaces also partake in radiative interactions, radiative flux uncertainties increase, not only because of inaccurate spatiotemporal distributions of their physical properties, but also because radiative properties of these elements have to be approximated and parameterized. All the above make the simulation of solar and thermal radiative processes a rather complex endeavor that burdens climate simulations with a substantial degree of uncertainty. Still, even before tackling radiative transfer involving clouds and aerosols, the more straightforward and well-defined problem of gaseous absorption needs to be evaluated and advanced. Unfortunately,

despite the significant share of CPU resources allocated to radiation in GCMs, the radiation codes in these models may still be inadequate in reproducing the radiative effects of increased greenhouse gases obtained by validated high-resolution codes. For example, a recent intercomparison [*Collins et al.*, 2006] of well-mixed greenhouse gas forcing calculations between line-by-line (LBL) RT codes and their speedier, but coarser, counterparts in GCMs used in the Intergovernmental Panel for Climate Change (IPCC) 4th Assessment Report, reported that for many of the cases analyzed, GCM codes exhibited "substantial discrepancies" relative to the detailed spectral LBL standards. These findings echoed earlier conclusions by the Intercomparison of Radiation Codes in Climate Models (ICRCCM) [*Ellingson and Fouquart*, 1991], indicating that progress in the intervening decade and a half was not as great as one would have hoped.

[3] While the synthetic cases employed in these previous studies had the benefit associated with well-defined controlled experiments, the baseline reference calculations themselves were not validated. This shortcoming was recognized by the organizers of the past intercomparisons who recommended the remedy of comprehensive field experiments that would measure the relevant atmospheric parameters and spectrally resolved radiation [*Fouquart et al.*, 1991; *Ellingson and Wiscombe*, 1996]. Such capabilities are now readily available with multi-instrument observational programs such as the Atmospheric Radiation Measurement (ARM) program in the U.S. [*Stokes and Schwartz*, 1994], now part of the Atmospheric System Research (ASR) program, and the European Cloudnet network [*Illingworth et al.*, 2007]. Real-world conditions at these observational sites include, however, the effects of spatially variable cloud, aerosol, and surface reflectance, and therefore pose significantly greater challenges when pursuing spectral or even broadband agreement between simulated and observed radiative fluxes. On the other hand, such observations, when well characterized, provide the opportunity to evaluate GCM radiation codes under a variety of conditions, a fact exploited by the intercomparison effort described in this paper, the Continual Intercomparison of Radiation Codes (CIRC).

[4] In an earlier paper *Oreopoulos and Mlawer* [2010] explained that CIRC was conceived as an evolving and regularly updated permanent reference source for evaluation of GCM-class RT codes that would facilitate their steady improvement. CIRC seeks to become the standard against which RT code performance is documented in scientific publications and coordinated joint modeling activities such as GCM intercomparisons. The LBL model that provides the reference calculations for CIRC has been validated against quality high-resolution spectral measurements from surface-, satellite-, and aircraft-based instruments [e.g., *Shephard et al.*, 2009; *Delamere et al.*, 2010], and represents the state of the science. Therefore, good agreement between the calculations of a particular GCM RT code with corresponding LBL calculations can be considered an indication of the quality of the faster code. While it is understood that CIRC reference calculations at any time would reflect spectroscopic knowledge that may not be perfect, by

keeping CIRC up to date with algorithmic and spectral database improvements as they become available, and by gradually expanding the effort with new cases, a valuable service to the radiation modeling community will be provided for years to come.

[5] This paper presents results from Phase I of CIRC, designed to test RT codes under relatively noncomplex atmospheric conditions, i.e., either cloudless skies or skies fully covered by homogeneous liquid clouds. Submissions from 13 solar and 11 thermal infrared codes are analyzed against reference LBL calculations. Besides overall performance, we also delve into particular aspects of RT code behavior exposed by unique features in the specification of individual cases. In addition, simplified versions of the cases help us isolate various sources of error.

2. The CIRC Phase I Data Set

2.1. The Cases

[6] Table 1 provides a summary of the seven baseline (primary) observation-based cases used in Phase I to test RT code performance. Two of the seven cases (Cases 6 and 7) include overcast liquid phase clouds that have very different condensate amounts. The cloudy cases were selected for their apparent homogeneity as indicated by low temporal variability of the shortwave (SW) downwelling fluxes at the surface measured by pyranometers. The high liquid water path (LWP) cloud case (Case 6) was observed at ARM's Southern Great Plains (SGP) observation site, while the cloud of modest LWP (Case 7) was observed during the 2005 deployment of the ARM Mobile Facility in Pt. Reyes, CA (PYE). The remaining five cases are cloudless. Three were observed at SGP (Cases 1–3), and another at ARM's Northern Slope of Alaska (NSA) site (Case 4). The remaining baseline cloudless case (Case 5) is partly synthetic; it is exactly identical to Case 4 except for the concentration of CO_2, which was doubled from 375 ppm to 750 ppm. The cloudless cases were selected to represent a variety of humidity, aerosol, solar zenith angle and surface albedo conditions. All six observed cases were selected based on their good broadband radiative closure (agreement between measured and calculated fluxes) at both the surface (SFC) and top of the atmosphere (TOA). Except for Case 7, the initial assessment for broadband radiative flux (irradiance) closure came from ARM's Broadband Heating Rate Profile (BBHRP) evaluation product (see Appendix A) [*Mlawer et al.*, 2002]. Radiative closure was subsequently verified with more accurate (LBL) calculations (see Appendix A) of which broadband results are shown in Table 1. In the longwave (LW) part of the spectrum, flux closure was also confirmed spectrally by comparing downwelling surface radiance measurements from the upward pointing Atmospheric Emitted Radiance Interferometer (AERI) [*Knuteson et al.*, 2004] (see Appendix A) with LBL calculations following *Turner et al.* [2004] (see example in the work of *Oreopoulos and Mlawer* [2010]). The difference between LW measurements and observations was also expressed as fluxes within the bands of the RRTM-LW [*Mlawer et al.*, 1997] radiation code by performing the following steps: The radiances and fluxes of the LBL code

Table 1. Synopsis of the Seven CIRC Phase I Primary (Baseline) Cases[a]

Date (Site)	Case	SZA	PWV (cm)	τ_{aer}[b]	LWP (gm^{-2})	LW_{SFC}[c]	LW_{TOA}[c]	SW_{SFC}[c]	SW_{TOA}[c]
25 September 2000 (SGP)	1	47.9°	1.23	0.04		289.7	301.7	705.9	169.8
						288.2	**304.3**	**701.2**	**175.0**
19 July 2000 (SGP)	2	64.6°	4.85	0.18		441.8	288.6	345.4	127.8
						439.3	**292.6**	**348.0**	**117.1**
4 May 2000 (SGP)	3	40.6°	2.31	0.09		336.4	277.6	772.5	159.6
						333.0	**280.8**	**773.1**	**173.6**
3 May 2004 (NSA)	4	55.1°	0.32	0.13		194.7	229.1	638.9	425.8
						192.4	**230.5**	**642.8**	**422.9**
3 May 2004 (NSA, 2 × CO$_2$)	5[d]	55.1°	0.32	0.13		**N/A**	**N/A**	**N/A**	**N/A**
						195.7	**229.2**	**641.3**	**422.7**
17 March 2000 (SGP)	6	45.5°	1.90	0.24	263.4	339.0	234.8	97.6	623.2
						335.2	**241.8**	**92.1**	**628.8**
6 July 2005 (PYE)	7	41.2°	2.42		39.1	373.2	284.0	479.8	356.0
						372.6	**280.2**	**473.7**	**356.4**

[a]SGP, Southern Great Plains; NSA, North Slope of Alaska; SZA, Solar Zenith Angle; PWV, Precipitable Water Vapor; LWP, Liquid Water Path.

[b]The aerosol optical depth (τ_{aer}) is for 0.55 μm.

[c]Columns for LW_{SFC}, LW_{TOA}, SW_{SFC}, and SW_{TOA} show observed and line-by-line (LBL)-calculated (in bold) flux values (in Wm^{-2}) at the surface (SFC, downwelling) and the top of the atmosphere (TOA, upwelling) for both the thermal/longwave (LW) and solar/shortwave (SW) parts of the spectrum. Observed TOA fluxes are from GOES using narrowband-to-broadband conversion algorithms or from CERES (Case 4), while observed SFC fluxes come from ARM instruments.

[d]Case 5 is as Case 4 but with doubled CO$_2$.

(described in Appendix A) were separately added within each RRTM band and the ratio of these sums was calculated; the ratio was then applied to the AERI-LBL radiance residuals of each RRTM band to obtain flux residuals that were subsequently added to obtain broadband flux residuals. The project webpage, http://circ.gsfc.nasa.gov/CIRC_cases. html shows these results and accompanying plots. Further details on the construction of the cases are provided in Appendix A.

[7] With the intercomparison underway, some CIRC participants suggested that the interpretation of model performance would be aided by including additional "subcases" that are simplified variants of the above seven baseline cases. These supplemental cases with simpler atmospheric and surface specifications are, of course, no longer radiatively constrained by observations. The subcases were constructed by imposing one or more of the following simplifications: (1) spectrally invariant SW albedo, (2) no aerosol, and (3) no cloud. A complete list of the subcases is provided in Table 2. A total of 16 SW subcases and 2 LW subcases resulted after applying various permutations of these simplifications. Changes in surface albedo and aerosol only affected SW cases since the LW cases ignore aerosols and assume a surface emissivity of unity across the full LW spectral range.

2.2. CIRC Data Sets

[8] In order to build the intercomparison cases, assure their robustness, and provide a set of RT reference calculations, several steps involving a variety of data sets are necessary. Much of this information can be found on the project website, http://circ.gsfc.nasa.gov. For completeness, we also include most of it here in Appendix A. The input for six of the seven cases (the exception being Case 7) is based on the ARM BBHRP evaluation data set, with details on how specific cases were built included in section A1 of Appendix A. Sections A2 and A3 provide details about the reference LBL RT calculations and the sources and treatment of the radiative observations used to ensure that a

satisfactory degree of radiative closure was achieved for all cases.

3. Requested Submissions and Participation

[9] Participants were asked to submit broadband irradiances (fluxes) at the atmospheric column boundaries (TOA and SFC) and heating rate profiles for the same vertical grid as in the provided atmospheric profiles. Specifically, the fluxes requested were: for LW, the upwelling flux at TOA and the downwelling flux at SFC; for SW, the upwelling flux at TOA, the upwelling flux at SFC, and the downwelling flux at SFC, both total and direct. It was requested that the latter be (if possible) the direct flux at the surface without contributions of forward scattered radiation along the direction of the solar beam. No spectral results

Table 2. CIRC Phase I Shortwave and Longwave Subcases

Subcase[a]	Description[b]
SW 1a	As SW Case 1, but with $R_{sfc} = 0.196$
SW 1b	As SW Case 1(a), but with no aerosol
SW 2a	As SW Case 2, but with $R_{sfc} = 0.188$
SW 2b	As SW Case 2(a), but with no aerosol
SW 3a	As SW Case 3, but with $R_{sfc} = 0.171$
SW 3b	As SW Case 3(a), but with no aerosol
SW 4a	As SW Case 4, but with $R_{sfc} = 0.670$
SW 4b	As SW Case 4(a), but with no aerosol
SW 5a	As SW Case 5, but with $R_{sfc} = 0.670$
SW 5b	As SW Case 5(a), but with no aerosol
SW 6a	As SW Case 6, but with $R_{sfc} = 0.136$
SW 6b	As SW Case 6(a), but with no aerosol
SW 6c	As SW Case 6(a), but with no cloud
SW 6d	As SW Case 6(c), but with no aerosol
SW 7a	As SW Case 7, but with $R_{sfc} = 0.164$
SW 7b	As SW Case 7(a), but with no cloud
LW 6a	As LW Case 6, but with no cloud
LW 7a	As LW Case 7, but with no cloud

[a]SW, shortwave; LW, longwave.

[b]R_{sfc} is the spectrally invariant surface albedo obtained as the weighted average of the original spectral albedo and the reference incoming spectral irradiance at the surface.

Table 3. LW Codes Participating in CIRC Phase I

Model Index	Brief Model Description	In GCM?	Experiment Variants	Submitter	References
0	LBLRTM v.11.1– AER_V_2.0 (based on HITRAN 2004), MT_CKD_2.0, LBL	No	None	Delamere, Mlawer	Clough et al. [2005]
1	RRTM-LW, 10–3250 cm^{-1}, CKD, 16 bands, 256 g-points	No	None	Iacono, Mlawer	Mlawer et al. [1997]; Clough et al. [2005]
2	RRTMG-LW, 10–3250 cm^{-1}, CKD, 16 bands, 140 g-points	Yes	None	Iacono	Mlawer et al. [1997]; Iacono et al. [2008]
3	CLIRAD-LW, 0–3000 cm^{-1}, k-distribution and one-parameter scaling, 10 bands, 85/113 k-points	Yes	High/low accuracy	Oreopoulos	Chou et al. [2001]
4	CCC 0–2500 cm^{-1}, CKD, 9 bands, 56 g-points	Yes	With/without scattering	Cole, Li	Li [2002]; [Li and Barker, 2002, 2005]
5	FLBLM, 40–3000 cm^{-1}, LBL,	No	None	Fomin	Fomin [2006]
6	FKDM, 40–3000 cm^{-1}, CKD, 23 g-points	No	None	Fomin	Fomin [2004]
7	CAM 3.1, 0–2000 cm^{-1}, absorptivity-emissivity approach	Yes	Treatment of lowest level air temperature	Oreopoulos	Collins et al. [2002, 2004]
8	FLCKKR (LW), 0–2200 cm^{-1}, CKD, 12 bands, 67 g-points	No	None	Rose	Fu and Liou [1992]; Fu et al. [1997]
9	RRTMG-LW (as implemented in FMI/ECHAM5.4), 10–3000 cm^{-1}, 16 bands, 140 g-points	Yes	None	Räisänen	Mlawer et al. [1997]; Iacono et al. [2008]
10	ES, 1–2995 cm^{-1}, 9 bands/33 g-points, ESF of band transmissions	Yes	With/without scattering	Manners	Edwards and Slingo [1996]; Edwards [1996]
11	NASA-GISS, 50–2000 cm^{-1}, CKD, 33 g-points	Yes	With/without scattering	Zhang, Rossow	Zhang et al. [2004]

were requested. The spectral range of LW and SW calculations was left to the participants, but under the condition that the radiation source was solely the Earth's surface and atmosphere for the LW, and the Sun for the SW.

[10] Some of the input information provided to conduct the CIRC runs is typically not available in an operational GCM environment, for example, spectral surface albedo. On the other hand, the available input may be incomplete for some RT codes, e.g., those requiring separate albedos for the ground and an overlying vegetation canopy. While we wanted the submissions to come from runs where the model uses as much of the information provided as possible, it was recognized that this may have required modifications of the RT codes from their standard operational configuration. Therefore, submissions where the algorithms operated with assumptions and input more closely resembling default operational configuration were acceptable. For example, allowing for cloud scattering in the LW (even though there is no scattering in the reference LBL calculations) was not prohibited.

[11] Submissions from 11 LW and 13 SW RT codes (also referred to as "models" throughout the paper) were received. The codes are identified in Tables 3 and 4, along with some pertinent information, such as whether the code is currently implemented in a large-scale model (LSM), and the party responsible for the submission. For some of the codes multiple submissions were received, corresponding to different configurations of the runs (e.g., high- or low-accuracy modes, with and without LW cloud scattering, different weightings of the spectral surface albedo function, different treatment of cloud optical properties, etc.). Some of these submission variants are used in the analysis presented below. Our primary comparisons use submissions corresponding to the configuration closest to the setup and assumptions of the LBL calculations. For example, when available, we use the LW simulations without cloud scattering, and the SW simulations where the band-averaged surface albedos were

calculated using the weighting by the spectral irradiance reaching the surface. Whenever a code's set up is fundamentally different than that of its counterparts, we make an effort to point out the differences in the presentation. While most participants submitted the full suite of requested calculations, some submissions were incomplete and are identified accordingly when affecting the comparison statistics.

4. General Assessment of Model Performance

4.1. Errors per Case, Model, and Flux Quantity

[12] A comprehensive representation of the performance of the participating codes is provided in Figures 1 (LW) and 2 (SW) which show a "map" of the percentage errors for each case n and participating model i (the index of each code comes from Tables 3 and 4) for the various flux types j. This percentage error is defined as:

$$e_{n,i,j}(\%) = 100 \times \frac{F_{n,i}^j - F_{n,0}^j}{|F_{n,0}^j|} \quad (1)$$

$F_{n,0}^j$ is the flux (irradiance) calculated with the reference LBL radiation code (LBLRTM for LW and LBLRTM/CHARTS for SW) for case n and flux type j, and $F_{n,i}^j$ is the corresponding flux from RT model i. The flux types for LW are the upwelling radiation at the TOA, the downwelling at the SFC and the net flux divergence of the atmospheric column defined as the net (down minus up) at TOA minus net at SFC. For the SW, the flux types are as above, with the addition of the diffuse component of the downward flux at the SFC defined as the total flux minus the direct horizontal flux. It should be noted that for the SW, the net flux divergence corresponds to the flux absorbed by the atmospheric column.

[13] In the LW, the best overall simulated flux type appears to be the downward flux at the SFC, but has two

Table 4. SW Codes Participating in CIRC Phase I

Model Index	Brief Model Description	In GCM?	Experiment Variants	Submitter	References
0	CHARTS v.4.04/LBLRTM v.11.1– AER_V_2.0 (based on HITRAN2004) MT_CKD_2.0, LBL	No	None	Delamere, Mlawer	*Moncet and Clough* [1997]; *Clough et al.* [2005]
1	RRTM-SW, 0.2–12.2 μm, CKD, 14 bands, 224 g-points	No	None	Iacono, Mlawer	*Clough et al.* [2005]
2	RRTMG-SW, 0.2–12.2 μm, CKD, 14 bands, 112 g-points	Yes	None	Iacono, Mlawer	*Iacono et al.* [2008]
3	CLIRAD-SW, 0.175–10 μm, 11 bands, pseudomonochromatic/k-distribution hybrid, 38 g-points	Yes	Two R_{sfc} averaging methods	Oreopoulos	*Chou et al.* [1998]; *Chou and Suarez* [1999]
4	CCC, 0.2–9.1 μm, CKD, 4 bands, 40 g-points	Yes	Three R_{sfc} averaging methods	Cole, Li	*Li and Barker* [2005]; *Li et al.* [2005]
5	FLBLM/HITRAN 11v, 0.2–10 μm, LBL	No	None	Fomin	*Fomin and Mazin* [1998]
6	FKDM, 0.2–10 μm, CKD, 15 g-points	No	Two treatments of cloud optical properties	Fomin	*Fomin and Correa* [2005]
7	CAM 3.1, 0.2–5.0 μm, 19 spectral and pseudospectral intervals	Yes	Two R_{sfc} averaging methods	Oreopoulos	*Briegleb* (1992); *Collins* [2001]; *Collins et al.* [2002, 2004]
8	FLCKKR (SW), 0.2–4.0 μm, CKD, 6 bands, 54 g-points	No	Two R_{sfc} averaging methods	Rose	*Fu and Liou* [1992]
9	FMI/ECHAM5.4, 0.185–4 μm, 6 bands, Padé approximants to fit transmission functions	Yes	Two R_{sfc} averaging methods	Räisänen	*Fouquart and Bonnel* [1980]; *Cagnazzo et al.* [2007]
10	Edwards-Slingo 0.2–10 μm, 6 bands/20 g-points, ESF of band transmissions	Yes	Two R_{sfc} averaging methods	Manners	*Edwards and Slingo* [1996]
11	NASA-GISS v. D, 0.2–5.0 μm, CKD, 15 g-points	Yes	Three R_{sfc} averaging methods	Zhang, Rossow	*Zhang et al.* [2004]
12	COART, 0.25–4.0 μm, 26 bands, k-distribution	No	None	Jin	*Jin et al.* [2006]
13	CLIRAD-SW modified, 8 bands, CKD, 15 k-points	No	Two R_{sfc} averaging methods	Oreopoulos, using code by Tarasova	*Tarasova and Fomin* [2007]

instances of flux errors outside the ±3% range of the color bar (Cases 4 and 5 for Model 3). The TOA flux error is never outside the ±3% bounds. Model 11 exhibits a persistent overestimate of ~1.5% in the downward flux and an underestimate of ~2% in the TOA upward flux (including cloudy cases). Since the upward flux at the surface depends solely on surface temperature according to the Stefan-Boltzmann law and the downward LW flux at TOA is zero, the errors in net flux divergence (which relates to the total heating/cooling rate of the entire atmospheric column) can be smaller than the errors in the TOA up or SFC down fluxes when error cancellation occurs. For about 70% of model-case combinations the net flux divergence indeed exhibits smaller errors than either of the TOA and SFC flux errors. But from those models that maintain SFC and TOA flux errors within ±1% of LBL, namely, Models 1, 2, 4, 9, and 10 (and 8 if cloudy Case 6 is excluded), only Models 1, 2, and 4 remain within ±1% of LBL once net flux divergence is also taken into account, i.e., the cancellation works mostly for the benefit of the poorer performing models. Note that Model 5, an LBL class model, belongs to neither of these two model groups.

[14] The boundaries of the color bars need to be significantly expanded to accommodate the percentage errors in the SW (Figure 2). For the TOA up flux and the SFC down flux there seems to be a general tendency to overestimate the flux, i.e., most models are too transmissive and too reflective. It comes therefore as no surprise that very substantial underestimations of the flux absorbed by the atmospheric column are produced by many models. Also, with the exception of Models 1, 2, and 12 (and to some extent 8), the diffuse downwelling flux at the surface seems to be a quantity that the participating RT codes have difficulty calculating (or defining). Errors outside the color bar limits of ±10% are quite common, with Models 3, 4, 7, and 9 being the worst offenders. The participant representing Model 10 pointed out that a good estimate of the diffuse flux is not possible for his model with a single calculation because rescaling of the phase function is needed for more accurate total SW fluxes, while nonrescaled calculations are needed for more accurate direct SW fluxes; an accurate SW diffuse flux could therefore be obtained only by subtracting two separate calculations. The total downwelling flux at the SFC for the thick cloud case (Case 6) also poses difficulties for some models: Models 3, 4, 5, 6, 6a (a variant of Model 6 with different treatment of cloud optical properties), 9, and 13 exhibit substantial underestimates of the downwelling flux. Overall, the SW errors depicted in Figure 2 are quite disheartening. To some extent, they are attributable to the nongaseous scatterers embedded within our test atmospheres, as shown in Figure 5d (discussed later). If one considers only the "pristine" (i.e., cloud- and aerosol-free) cases with spectrally constant surface albedo (Subcases 1b, 2b, 3b, 4b, 5b, 6d, and 7b), which are designed to test the treatment of molecular absorption and scattering, the TOA up and SFC down fluxes exhibit significantly better agreement with the corresponding reference calculations (with the exception of Model 11, Case 2b, the errors are always within ±5%). On the other hand, in the absence of clouds and aerosols, the underestimation of gaseous absorption is
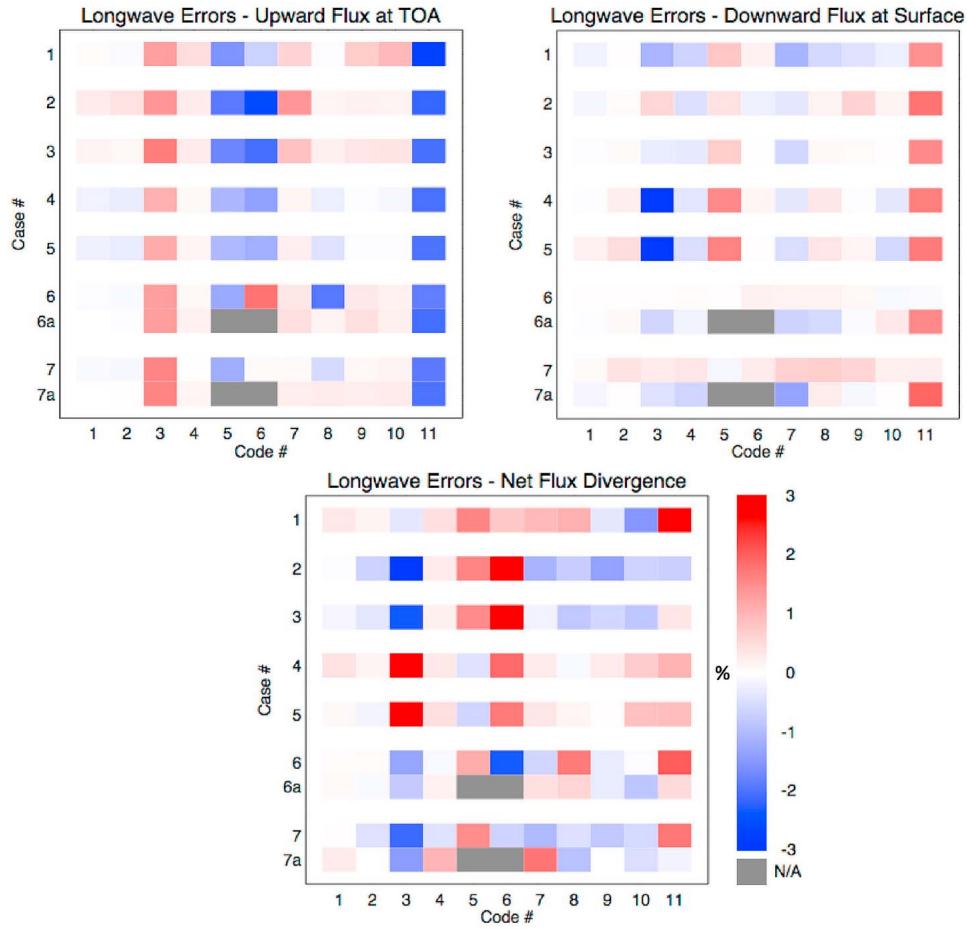
Figure 1. Percentage errors as defined by equation (1) of each participating model for each case for longwave (LW) upward flux at the top of the atmosphere (TOA), downward flux at surface (SFC), and net flux divergence. Gray indicates unavailability of submissions. Errors outside the color bar range are assigned the extreme colors of the color bar.

exposed more clearly, with absorbed flux errors becoming larger, especially for Cases 6 and 7. Only Models 2 and 5 (an LBL code) achieve absorption errors within ±2.5% for all the pristine cases.

4.2. Overall Model Errors per Flux Quantity

[15] A more compact view of model performance can be obtained by averaging errors over all cases, and these are shown in Figures 3 and 4. Specifically, for each model i and flux type j we show the mean error (bias) and the root mean square (rmse) of all cases, defined respectively as:

$$\overline{e_i^j} = \frac{\sum\limits_{n=1}^{N_c} \left(F_{n,i}^j - F_{n,0}^j \right)}{N_c} \tag{2}$$

$$rmse_i^j = \sqrt{\frac{\sum\limits_{n=1}^{N_c} \left(F_{n,i}^j - F_{n,0}^j \right)^2}{N_c}} \tag{3}$$

where N_c is the number of cases (including subcases), i.e., 9 for LW and 23 for SW (7 and 21 for Model 6). All quantities are expressed in Wm^{-2}. Notwithstanding the

small number of CIRC Phase I cases, cancellations are allowed in the calculation of the mean error per equation (2) to acknowledge the fact that in an operating environment average performance on a wide range of atmospheric input is also highly relevant. On the other hand, the rmse is perhaps a superior metric of intrinsic model performance, since it accumulates the individual errors. For example, the fairly significant root mean squares of LW flux divergence for Models 7 and 8 and of the SW column absorption for Model 6 are clear indications that a good performance on average does not signify consistently good performance across the board. To facilitate bias and RMS error comparisons, we plot in Figures 3 and 4 the absolute magnitude of the bias, but distinguish negative values by overlaying a minus sign above the corresponding bars.

[16] Another benefit of this type of presentation is that it identifies the flux component best simulated by the participating models. For the LW, the downwelling SFC flux is simulated best; for the SW, the upwelling TOA flux is slightly better simulated than the other fluxes. The alarmingly large percentage errors of some models for the absorbed SW flux and diffuse downwelling flux shown previously in Figure 2 are reaffirmed here with mean errors above 10 Wm^{-2}. Note that for column absorption the
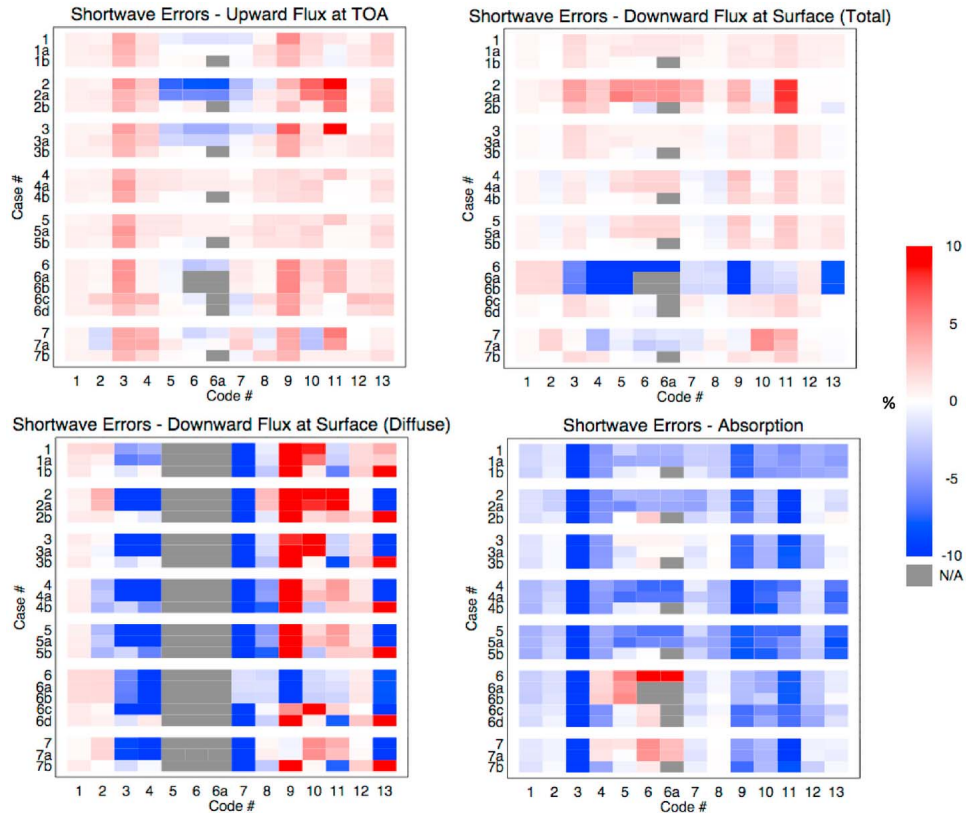
Figure 2. Percentage errors as defined by equation (1) of each participating model for each case for shortwave (SW) upward flux at TOA, total downward flux at surface, diffuse downward flux at surface (difference between total flux and the direct solar flux), and absorption. Gray indicates unavailable submissions. Errors outside the color bar range are assigned the extreme colors of the color bar. Model 6a is a variant of Model 6 with optical properties of the scatterers described in greater spectral detail within the k distribution bands following *Fomin and Correa* [2005].

numerical values of mean error and rmse are almost identical for most models (except Models 5 and 6), reflecting the previously identified near-universal underestimate of absorption.

4.3. Overall Model Errors for All Flux Types

[17] Even more compact error measures that combine the different flux components can be used to assess individual model performance. For each participating model i, one can define a total percentage error that accounts for different flux types and cases as follows:

$$e_{tot,i}(\%) = \frac{100}{N_c N_f} \sum_{n=1}^{N_c} \sum_{j=1}^{N_f} \frac{\left|F_{n,i}^j - F_{n,0}^j\right|}{\left|F_{n,0}^j\right|} \qquad (4)$$

where N_f is the number of flux types. For this analysis we exclude the diffuse SW surface flux which is not only simulated poorly, but is also unavailable for Models 5 and 6, so $N_f = 3$ for both the SW and LW. In addition to calculating this error for all 9 LW and 23 SW cases, we also perform a second calculation for only the subcases of pure molecular atmospheres, i.e., excluding Cases 6 and 7 for the LW and excluding all but Subcases 1b, 2b, 3b, 4b, 5b, 6d and 7b for the SW.

[18] Figures 5a and 5b show the errors calculated using equation (4). Such an error metric could potentially serve as a simple way to rank model performance and to set thresholds of acceptable RT code performance for particular applications. Figures 5c and 5d reveal whether the errors calculated from equation (4) are smaller or larger when only simplified atmospheres and surfaces are considered. Given that in the LW the simplified atmospheres differ only in that nonscattering clouds have been removed, it is not surprising that performance does not in general improve in their absence. (Model 8, which performs better for clear skies, did not provide LW calculations without scattering for the cloudy cases). On the other hand, all models, with one exception, improve their performance in the SW for the simplified cloudless atmospheres with no aerosols and no spectral variation of surface albedo. The exception, Model 12, suffers primarily from the fact that its absorbed flux for the pristine cases is sufficiently inaccurate to push the error metric of equation (4) above the values corresponding to the full basket of cases.

4.4. LW Heating Rate Errors

[19] Heating (cooling) rate errors are only calculated for the LW since no reference LBL SW heating rate profiles are

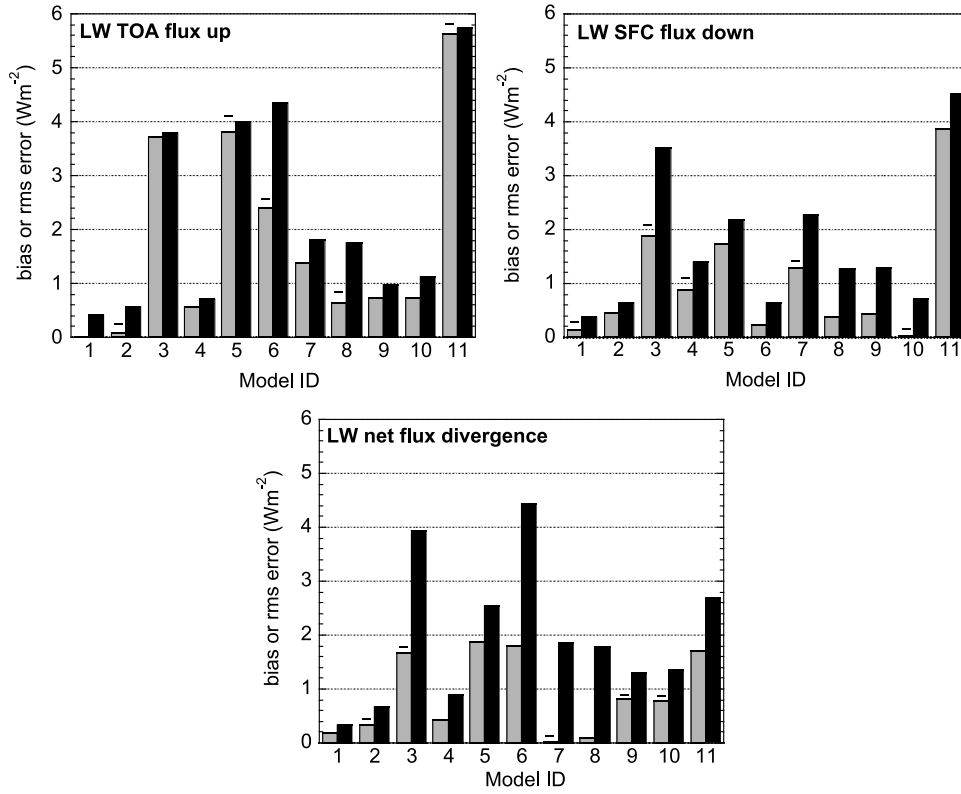**Figure 3.** Absolute value of mean (bias) errors in Wm^{-2} from equation (2) (gray bars) of the three flux types of Figure 1 over all cases, with root mean square errors from equation (3) (black bars). The mean errors that are negative (underestimates compared to line by line (LBL)) are identified by a minus sign above the bar.

available. We use only the five original clear-sky cases and the two cloudy cases in order to put Models 5 and 6, which did not submit results for the two LW subcases, on equal footing with the other models. The additional vertical dimension makes the evaluation of heating rate profile errors somewhat more challenging than for the column boundary fluxes examined so far. We settled on using a mass-weighted [e.g., *Räisänen and Barker*, 2004] heating rate (HR) root mean square error (rmse) for each model i calculated as follows:

$$rmse_i(HR) = \sqrt{\frac{\sum_{n=1}^{N_c} \sum_l \left(HR_{n,i}^l - HR_{n,0}^l\right)^2 \Delta p_n^l}{\sum_{n=1}^{N_c} \sum_l \Delta p_n^l}} \quad (5)$$

where the heating rate $HR_{n,i}^l$ (cooling rate, when a negative value is obtained) of model i in layer l for case n is given in (K/day) by:

$$HR_{n,i}^l = 86400(s/day) \times \frac{g}{c_p} \frac{\Delta F_{n,i}^l}{\Delta p_n^l} \quad (6)$$

In equation (6) Δp_n^l is the pressure thickness of layer l for case n, c_p is the specific heat of air at constant pressure, g is the gravitational constant, and $\Delta F_{n,i}^l$ is the flux divergence of Model i in layer l of case n. As before, the index 0 indicates the reference LBL model. For the clear-sky cases we calculate the

HR rmse separately for the parts of the atmosphere below and above 200 hPa (a proxy separator between troposphere and stratosphere also used by *Collins et al.* [2006]). For the cloudy cases we calculate the rmse only for the layers below 200 hPa. Figure 6 shows the results. Note that for Model 5 we used the submission with scattering included since the submission with no scattering was at a higher vertical resolution and could thus not be used in the heating rate comparison.

[20] Models 6 and 7 stand out for their poor simulation of HRs below 200 hPa for cloudy atmospheres. Only one other model (Model 9) exceeds 1 K/day in HR rmse for the cloudy cases. Six models simulate tropospheric heating rates more accurately and five simulate stratospheric heating rates better. If one seeks to correlate performance in column boundary fluxes with performance in HRs, a relationship is indeed apparent. The models underperforming in Figure 5a (Models 3, 5, 6, and 11) also do the same in terms of HR rmse. Models 7 and 8 (the latter including scattering in the cloudy cases), the next two with the largest percentage error (from equation (4)) in Figure 5, also exhibit substantial disagreements with the LBL results in their HR simulations.

5. Analysis of Possible Sources of Model Errors

5.1. Effects of Averaging Spectral Surface Albedo

[21] The surface albedo is a major contributor to the reflected solar flux at TOA, especially in cloudless conditions.
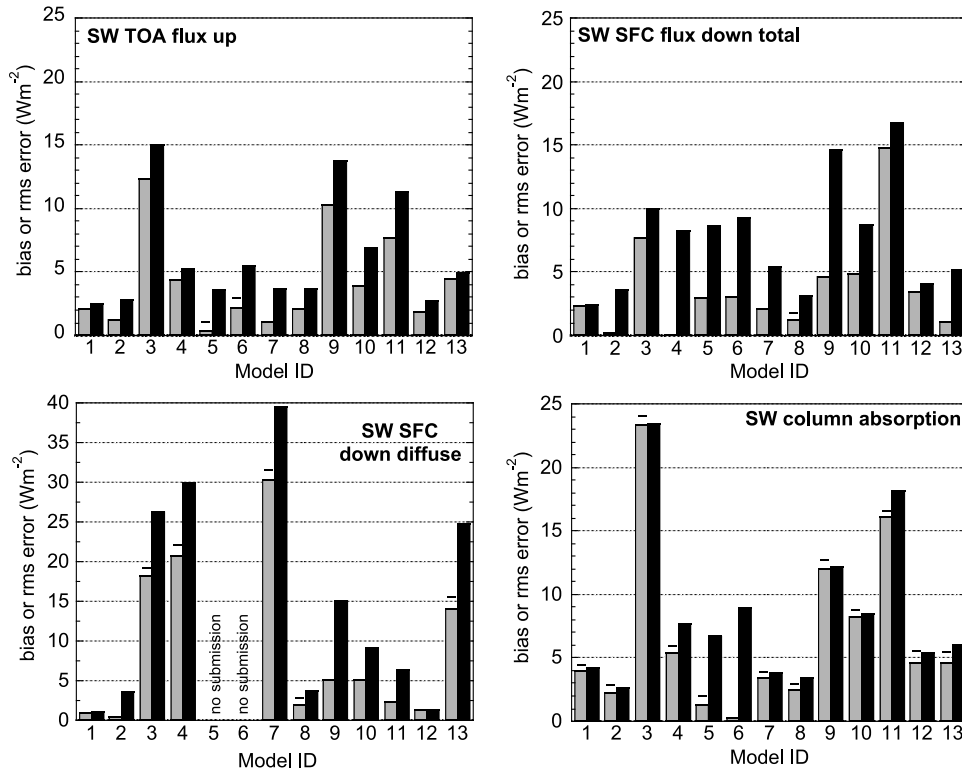
Figure 4. As in Figure 3 but for the four flux types of Figure 2.

While we provide the detailed spectral variation of surface albedo at 1 cm^{-1} resolution, virtually all participating models resolve the albedo into much coarser spectral bands. A number of models have only two bands, one for ultraviolet visible wavelengths and one for all longer solar wavelengths. It is therefore worth examining whether the specifics of how the detailed spectral surface albedo is coarsened by the various models have a substantial impact on the simulated TOA fluxes. Specifically, the participating models are examined in terms of (1) the flux change when a spectrally variable surface albedo is used in place of a spectrally flat albedo, and (2) the difference in flux between two different ways of averaging the spectrally detailed albedo into coarse spectral intervals.

[22] To gain insight into the impacts of ignoring altogether surface albedo spectral variations we calculate the TOA flux differences between the baseline and "a" subcases (i.e., the difference between cases with spectrally resolved albedo and spectrally flat albedo) and average them separately for the clear (Cases 1–5) and cloudy cases (Cases 6–7), i.e.,

$$\overline{\Delta F_i^{TOA}} = \frac{\sum\limits_{n=1}^{N_c} \left(F_{n,i}^{TOA} - F_{n(a),i}^{TOA} \right)}{N_c} \tag{7}$$

where $N_c = 5$ for clear and $N_c = 2$ for cloudy. Figure 7 shows these average differences, with the LBL results shown as dashed lines.

[23] To examine the impact of spectrally coarsening the surface albedo, we estimate for Cases 1–4 the TOA flux

difference between calculations where the coarse-band albedo comes from weighting the spectral surface albedo values with the LBL spectral downwelling flux reaching the surface (these calculations were used for obtaining $F_{n,i}^{TOA}$ in equation (7)) and calculations where the incoming spectral solar radiation at the TOA was instead used for the weighting. Figure 8 shows these differences (positive indicates that the flux from surface flux weighting is larger) for those models that submitted results for both surface albedo weighting options.

[24] The LBL results in Figure 7 indicate that the average effect of surface albedo spectral variations is ~1.5 Wm^{-2} for cloudless atmospheres and ~−0.4 Wm^{-2} for cloudy atmospheres. The sign of the TOA flux difference is negative for Case 4, with its ice-driven surface albedo, and the two cloudy cases, and positive for all other cases (including Case 5). Five models stand out in their disagreement with LBL on the effect of surface spectral albedo variations: Models 6, 7, and 8 for producing the wrong difference sign for clear cases and Models 9 and 11 for large overestimates of the difference. Three of these models (7, 8, 11) join Models 4 and 5 in producing substantial discrepancies for the cloudy cases as well.

[25] The models least sensitive to the surface albedo weighting method are 4, 9, and 10 (Figure 8). The first uses four bands while the other two use six bands to resolve the surface albedo. For the other models, which have greater dependence on the weighting approach, all except Model 11 (which also resolves surface albedo into six bands) use two
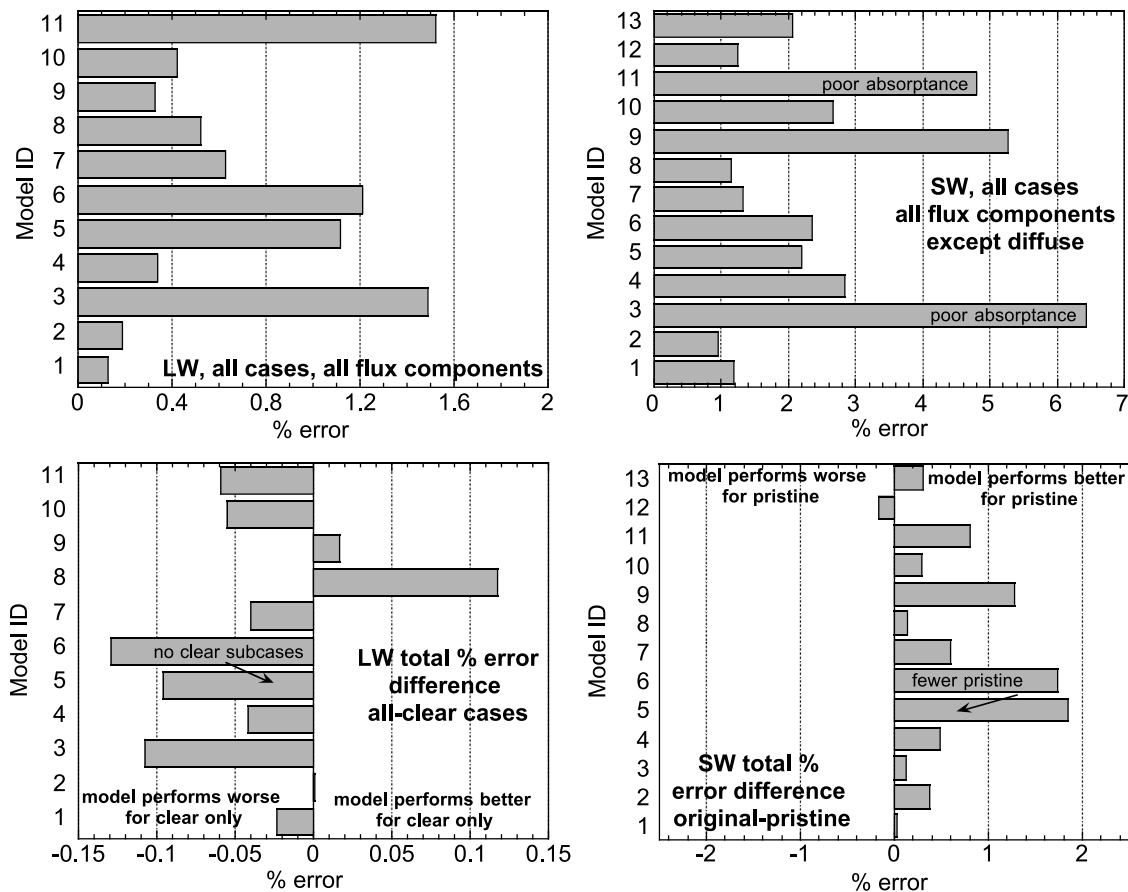
Figure 5. (a, b) The total error of each participating model according to equation (5) for all LW and SW cases. (c, d) The difference between total errors calculated for the full basket of cases (i.e., Figures 5a and 5b) and total errors calculated for the subset of the seven cloudless (clear) LW cases and the subset of seven pristine SW cases (no aerosol, spectral flat surface albedo) of Table 2.

bands, one for the ultraviolet/visible range and one for all remaining wavelengths above 0.7 μm. These results reflect the fact that, unlike the spectral albedo, a wideband average albedo is not an intrinsic surface property. In principle, different atmospheric conditions give rise to distinct wideband surface albedos because of changes in the spectral distribution of solar flux reaching the surface. This dependence on the overlying atmosphere should be smaller than shown in Figure 8 because the relative energy distributions at TOA and SFC differ more greatly than two SFC flux distributions under different atmospheric conditions. The greater differences in Figure 8 compared to Figure 7 underscore the importance of appropriate weighting of spectral albedos to obtain a wideband value. Ignoring all spectral variations in albedo leads to smaller errors at the TOA (Figure 7) compared to a calculation where band-specific albedo values were derived using an inappropriate (e.g., the incoming radiation at TOA) weighting (Figure 8), as long as the single broadband albedo value is derived from proper weighting. This result indicates that due consideration must be given to how wideband albedos are derived from highly resolved spectral albedo functions. The subtle effects of surface albedo averaging will reemerge again in the CO_2 forcing analysis of section 5.3.

5.2. Cloud and Aerosol Radiative Effect Errors

[26] Cloud and aerosol (SW-only) radiative effect errors by the participating models can be best isolated by using the spectrally constant surface albedo subcases, which are devoid of differences that may arise by the coupling of
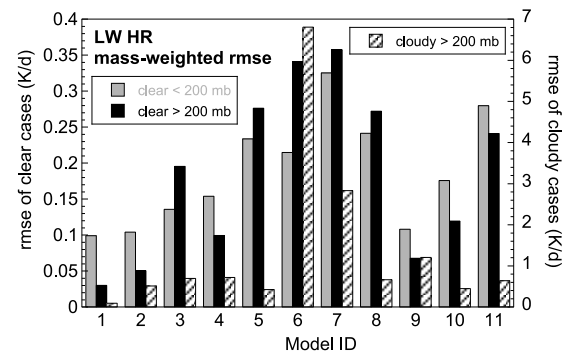

Figure 6. The mass-weighted LW heating rate (HR) root mean square error of each participating model derived from equation (5). Three HR root mean squares are calculated: for the original five clear cases below the 200 hPa level, for the original five clear cases above the 200 hPa level, and for the two cloudy cases below the 200 hPa level.
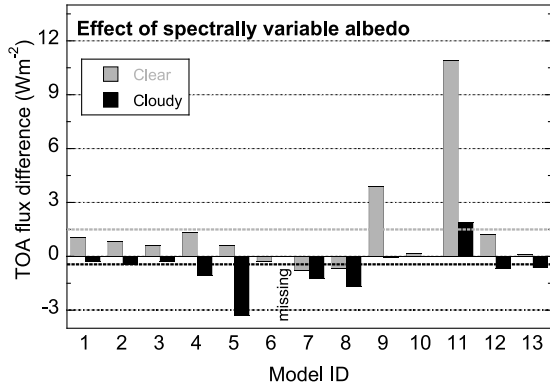
Figure 7. TOA flux difference between the baseline and "a" subcases (i.e., the difference between cases with spectrally resolved albedo and spectrally flat albedo) averaged separately for the clear (Cases 1–5) and cloudy cases (Cases 6–7) per equation (7). The gray and black dashed lines denote the clear and cloudy LBL results, respectively.

aerosol/cloud scattering with the coarsely averaged albedo. The error in LW and SW cloud radiative effect for model i and flux type j is therefore defined as:

$$e^j_{n(y)-n,i} = \left(F^j_{n(y),i} - F^j_{n,i} \right) - \left(F^j_{n(y),0} - F^j_{n,0} \right) \quad (8a)$$

$$e^j_{n(y)-n(x),i} = \left(F^j_{n(y),i} - F^j_{n(x),i} \right) - \left(F^j_{n(y),0} - F^j_{n(x),0} \right) \quad (8b)$$

where $n = 6$ or 7. For the LW $y =$ "a" in equation (8a). To evaluate the SW cloud radiative effect error for Case 6 we select the aerosol-free subcases, so $x =$ "b" and $y =$ "d" in equation (8b); for Case 7, which does not include aerosols, $x =$ "a" and $y =$ "b." These cloud radiative effect errors (both SW and LW) are shown in Figure 9.

[27] Because the clouds of Cases 6 and 7 are low, their LW radiative effects are small at TOA and large at the SFC. The reference TOA radiative effects are 6.5 Wm^{-2} for Case 6 and 5.9 Wm^{-2} for Case 7 (second term on the right hand side of equation (8a) where cloudy flux is subtracted from clear flux) while the reference SFC radiative effects are −65.1 Wm^{-2} and −53.3 Wm^{-2}. LW cloud radiative effect errors remain approximately within ±5 Wm^{-2} for all flux types for both cloudy cases, with the exception of Model 7 which is slightly outside this range for the downward SFC flux and column flux divergence, and are generally larger for the less optically thick cloud of Case 7. The percentage errors stay below 13%, with the single exception of Model 8, which exhibits huge TOA errors. It is important to note that the LW calculations of Model 8 include scattering, which is not accounted for in the reference LBL runs.

[28] SW cloud radiative effect errors are much larger than their LW counterparts and are generally of greater magnitude for the optically thicker cloud of Case 6. Due to smaller absolute values and occasional error cancellations, absorbed flux radiative effect errors can be small even if TOA and SFC radiative effect errors are large. Overall, it appears that the cloud radiative effect on the downwelling SFC flux is simulated slightly worse than the TOA radiative effect. In interpreting these results, however, one has to take into

account the much larger magnitude of the SW cloud radiative effect compared to its LW counterpart: the reference TOA radiative effect is −500.5 Wm^{-2} for Case 6 and −202.3 Wm^{-2} for Case 7, while the respective reference SFC radiative effects are 651.2 Wm^{-2} and 286.4 Wm^{-2}. Percentage errors are always below 10% for TOA and SFC, while absorption radiative effect errors can exceed 20% (Models 4, 5 and 11).

[29] Aerosol radiative effects have been investigated in CIRC only for the SW. Similar to equations (8a) and (8b), we calculate the flux differences between aerosol-free and aerosol-laden cases, subtract the corresponding LBL flux difference, and then average over the $N_c = 6$ cases containing aerosols (Case 7 is aerosol free) to estimate the mean errors:

$$\overline{e^j_i} = \frac{\sum\limits_{n=1}^{N_c} e^j_{n(y)-n(x),i}}{N_c} \quad (9)$$

where $x =$ "a," $y =$ "b" for Cases 1–5 and $x =$ "c," $y =$ "d" for Case 6. The SW aerosol radiative effect errors calculated from equation (9) are shown in Figure 10. Models 5 and 6 are noticeable for their inferior performance in the simulation of the SFC and column absorption radiative effects even though the TOA radiative effect is simulated very well. The aerosol radiative effects themselves are smaller than the cloud radiative effects, so it is not surprising that errors expressed in Wm^{-2} are also smaller. The fact that nearly all models have positive "COL" errors in Figure 10 indicates that aerosol absorption is generally underestimated by the models.

[30] In comparing cloud and aerosol SW radiative effect errors, a fundamental feature of the construction of Phase I cases and subcases should be borne in mind, namely, that the aerosol extinction, asymmetry factor and single scattering albedo were given explicitly. The lack of spectral variation of the latter two optical properties presumably makes their incorporation in the participating radiation codes straightforward. In contrast, for the cloudy cases only the effective radius profile was given; participants had to determine
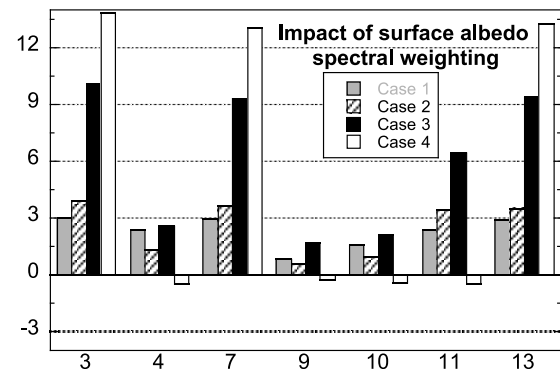

Figure 8. TOA flux difference for Cases 1–4 for select participating models between calculations where band-averaged albedo values were derived using as weights the values of reference LBL spectral downward flux at the surface, and calculations where the weights came from values of incident spectral flux at TOA. Models not shown did not provide data for both weighting schemes.
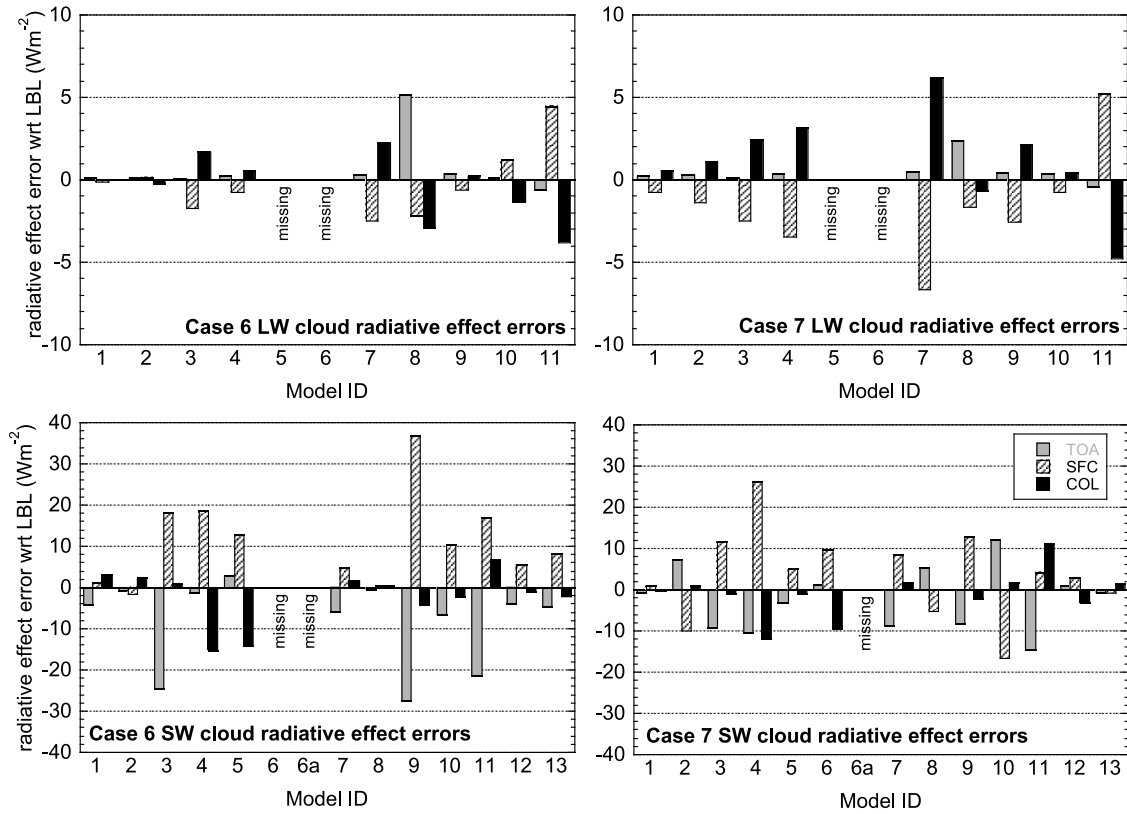
Figure 9. Cloud radiative effect errors for the (top) LW and (bottom) SW. The errors are defined as the radiative effect of the participating model minus the LBL radiative effect (see equation (8)). The radiative effects are defined as the difference of the cloudless fluxes minus the cloudy fluxes. The LW radiative effect is therefore calculated as the difference in fluxes between Subcase 6a and LW Case 6 and between Subcase 7a and LW Case 7. The SW cloud radiative effect is defined as the flux difference between SW Subcases 6d and 6b and SW Subcases 7b and 7a (spectrally invariant surface albedo and no aerosol). Three types of radiative effect are shown: for TOA flux up (TOA), for SFC flux down (SFC), and for the column flux divergence (COL; column absorbed flux in the SW).

band-wide spectral averages of cloud optical properties from this sole piece of information, as many GCM radiation codes do in an operational environment. If the detailed spectral variability of cloud optical properties used in the LBL calculations was instead provided for participants to properly average within their model's band structure, perhaps smaller cloud radiative effect errors would have resulted.

5.3. Carbon Dioxide Forcing Errors

[31] The CO_2 forcing can be defined as either:

$$f_{5-4,i}^{j} = F_{5,i}^{j} - F_{4,i}^{j} \tag{10a}$$

or:

$$f_{5(x)-4(x),i}^{j} = F_{5(x),i}^{j} - F_{4(x),i}^{j} \tag{10b}$$

Equation (10a) applies for both the LW and SW; equation (10b) applies only for the SW with either x = "a" or x = "b." In the SW therefore three CO_2 forcing calculations are possible, one that corresponds to the baseline case and two that correspond to the two subcases (spectrally flat albedo with aerosol and spectrally flat albedo with no
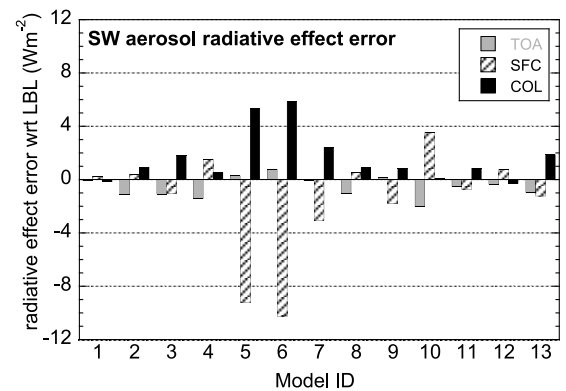

Figure 10. Aerosol radiative effect error (difference of participating model minus LBL radiative effect) in the SW. The radiative effect is defined as the difference of fluxes without aerosol minus fluxes with aerosol, specifically the difference between subcase "b" minus subcase "a" fluxes for Cases 1–5 and the difference between subcase "d" minus subcase "c" fluxes for Case 6 (there is no aerosol in Case 7). The radiative effect errors depicted are averages over all six cases (see equation (9)).
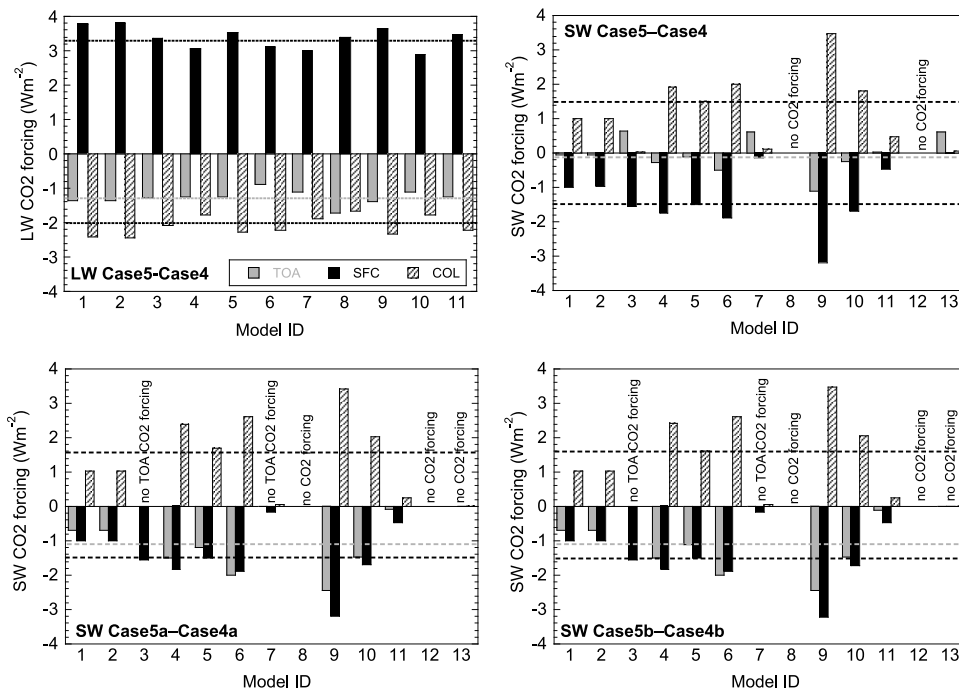
Figure 11. CO_2 forcing (flux difference between cases with 750 ppm CO_2 and 375 ppm CO_2) for LW and SW cases and subcases. The LBL forcings are indicated by the dashed lines. (a) For LW, the black line at a negative value is for the forcing in atmospheric flux divergence (COL), the gray line is for forcing of upwelling TOA flux, and the black line at a positive value is for the downwelling flux at SFC. (b–d) For SW, the lines at negative values are for TOA (gray) and SFC (black), and the line at positive values is for absorptance (COL). Further explanation is given in the text.

aerosol). We choose to show not the errors, but the forcings themselves in order to highlight an issue regarding the sign of the SW TOA forcing. They are shown in Figure 11, where the dashed lines depict the LBL reference results. In the LW, models divide in almost equal numbers to those that underestimate and those that overestimate the SFC and COL (flux divergence) forcing. For TOA forcing, however, only one model yields a notable overestimate (Model 8) while one (Model 6) exhibits a substantial underestimate. The magnitudes of the LW CO_2 forcing differ from LBL by typically less than 0.5 Wm^{-2}. Considering though the small absolute value of the forcing, percentage errors can be significant: the maximum deviation of 0.43 Wm^{-2} for TOA by Model 6 corresponds to a 32% error; the maximum deviation of 0.48 Wm^{-2} for SFC by Model 2 corresponds to a 14% error.

[32] When comparing the magnitude of the LW TOA CO_2 forcing with values frequently quoted in the literature for scenarios of doubled CO_2, one should keep in mind the following differences in our calculations: (1) CO_2 doubles from 375 ppm and not from preindustrial concentrations; (2) the atmosphere for which CO_2 is doubled is a cold and dry arctic atmosphere over a cold surface and not a standard, mean, or midlatitude atmosphere; and (3) our forcing does not refer to the tropopause, but the top of the atmospheric column.

[33] With regard to the SW results, Figure 11 indicates that the quality of CO_2 forcing estimates varies greatly. Concentrating on the CO_2 forcing of the spectrally flat albedo subcases shown in Figure 11c (the spectrally variable baseline cases have interpretation subtleties that will be discussed later), deviations from LBL calculations are quite large, as was previously found by *Collins et al.* [2006]. The poorest performer is Model 9 with percentage errors above 100%. For this model CO_2 doubling apparently implicitly increases the concentration of some of the other uniformly mixed absorbing gases, O_2, N_2O, CH_4 and CO. When the spectral response of this model to doubling CO_2 was further examined by the participant, flux changes were seen in spectral intervals where CO_2 is radiatively inactive. Such deficiencies can have significant unintended consequences if not eliminated from RT codes used in climate models.

[34] Whether forcing is calculated using the "a" and "b" subcases is irrelevant since CO_2 forcing occurs in the near-IR where aerosol radiative effects are minor (compare Figures 11c and 11d). The forcing of the baseline cases is, on the other hand, quite different from the forcing of the subcases. This is because of the effects of spectrally variable surface albedo. In the case of spectrally variable albedo, the LBL TOA forcing is minute, 0.14 Wm^{-2}, a consequence of the surface being very absorptive (i.e., low albedo) in the spectral regions in which CO_2 absorbs – essentially any radiation absorbed by the additional CO_2 would not have made it to the TOA anyway. For spectrally constant surface albedo the TOA forcing is 1 Wm^{-2} higher since the broadband spectral albedo was not adjusted to account for the slight shift in the relative spectral distribution of surface downwelling SW irradiance due to the increased absorption of the doubled CO_2. This result demonstrates that GCMs may commit fundamental errors in radiative forcing due to
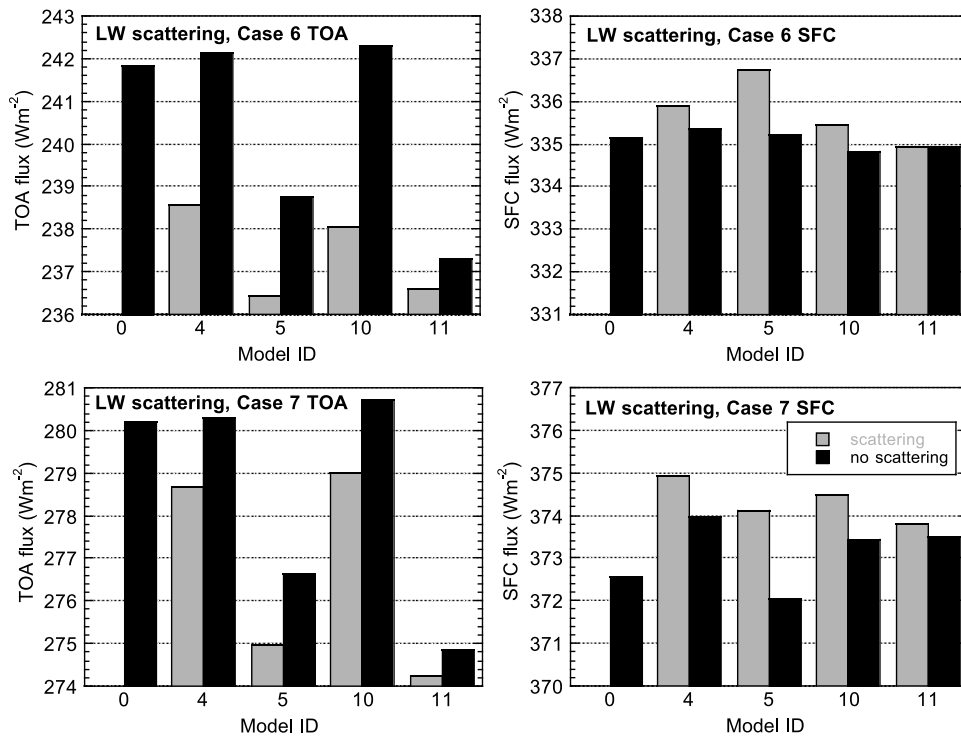
Figure 12. Comparison between LW fluxes at TOA and SFC with and without scattering for the participating models that submitted both types of calculations. The LBL reference fluxes also shown do not include scattering.

increases in greenhouse gas abundances if these subtle effects related to surface albedo are neglected. Two models, 3 and 7, designed to produce SW CO_2 forcing only at the SFC and not the TOA (see Figures 11c and 11d) end up producing positive TOA forcing (i.e., Case 5 reflecting more than Case 4 despite the larger CO_2 concentration) for this exact reason, a higher near-infrared surface albedo for Case 5 due to downwelling surface flux weightings that have changed from Case 4. This also the case for Model 13 which does not allow for CO_2 forcing neither at TOA or the SFC. This finding has been previously highlighted by *Oreopoulos and Mlawer* [2010]. Since in an operational GCM environment the effect of CO_2 on band-averaged surface albedo would most likely be neglected, a zero CO_2 TOA forcing would result for these two models, as intended by design.

5.4. Longwave Scattering

[35] The LBLRTM code does not account for cloud scattering in the LW and we therefore have no reference results to evaluate how accurately the impact of scattering is calculated by those participating models (4, 5, 10 and 11) that had the capability to provide fluxes both with and without scattering accounted for. A comparison between those models is shown in Figure 12; the no-scattering results from the LBL code are also included for reference. To facilitate the comparison of the differences between scattering and no-scattering results for both cloudy cases and the two fluxes at TOA and SFC, we chose a common range of 7 Wm^{-2} for all graphs. Scattering effects are in general greater at the TOA than at the SFC. Scattering affects the TOA upwelling

flux by reducing the transmission of upwelling radiation from within and below the cloud and by enhancing the upwelling flux due to backscattering of downwelling flux above the cloud. Figure 12 indicates that the former effect dominates (TOA upwelling is less with scattering), and this can be explained by the decrease of atmospheric temperature with height. The downwelling flux at the SFC is affected by backscattering of upwelling radiation below the cloud and reduction of downwelling flux from within and above the cloud. Figure 12 indicates that the former effect dominates (SFC downwelling is more with scattering), again because of the negative temperature lapse rate.

[36] Models 4 and 10 are quite similar with regard to their scattering effects, while Model 11 has the weakest scattering effect at the TOA and SFC for both cases. Model 5 has the largest scattering effects at the SFC, and exhibits the second weakest scattering impact (i.e., a lesser reduction in outgoing LW radiation) on the TOA flux of Case 6. The scattering effect on Case 6 TOA fluxes for Model 5 is, however, similar to that of Models 4 and 10. There is no obvious relationship between the impact of scattering in a particular model and how well its nonscattering calculation agrees with the LBL results. In general, the nonnegligible effects of cloud scattering in the LW depend on cloud optical thickness, cloud height and the temperature structure of the atmosphere, and their inclusion in LW RT codes will yield more accurate fluxes. For the CIRC Phase I cloudy cases, results from these limited submissions indicate effects as high as 4 Wm^{-2}. This value is comparable with the near-global (60°S–60°N) reduction of 3 W m^{-2} in the outgoing LW radiation due to

cloud scattering reported by *Costa and Shine* [2006], which amounts to ~10% of the global LW cloud radiative effect at the TOA.

6. Discussion

[37] The Continual Intercomparison of Radiation Codes (CIRC) is an ambitious effort to maintain a rigorous evaluation process for radiation codes used in variety of atmospheric and climate research applications. Radiation codes used in GCMs are of great interest because of the extensive use of this class of models in future climate prediction, but codes used in the calculation of broadband fluxes from retrieved geophysical parameters are also important to evaluate.

[38] This paper presents an analysis of performance for 11 thermal and 13 solar radiation codes that submitted results for the relatively simple clear and cloudy cases of CIRC Phase I. The reference LBL results used for the evaluation of submissions are publicly available for download at the CIRC website http://circ.gsfc.nasa.gov. Provided that CIRC participants will consent in having their submissions posted on the CIRC website, any interested individual will be able to perform their own code evaluation and examine aspects of code performance not covered in this paper.

[39] For our collection of cases, we found that errors in SW simulations, which have more degrees of freedom and parameters to specify, were larger than LW errors, a result previously found also by ICRCCM. Diffuse and absorbed SW fluxes are particular areas of concern. Obtaining the correct breakdown of total into direct and diffuse may be important for the simulation of surface processes such as vegetation growth and the carbon cycle in climate models with such capabilities. Previously found underestimates of SW absorption by less spectrally detailed models [*Ackerman et al.*, 2003] were confirmed again here. Another finding was that the number of bands available to resolve spectral surface albedo, as well as other details of albedo wideband averaging, can significantly alter TOA fluxes. Therefore, it may be beneficial to pay more attention to proper representations of spectral albedo variations. While LW fluxes were quite adequately simulated at the atmospheric column boundaries, net flux divergences within the atmosphere that determine heating (or cooling) rates were less accurate. Moreover, inclusion of cloud scattering in the LW, with potential impacts on the TOA and SFC fluxes of a few Wm^{-2} will have a nonnegligible positive impact on RT code performance. In general, cloud radiative effect errors for both the LW and SW stayed below 10% for the vast majority of participating codes, a result which is quite satisfactory considering that with no explicit cloud scattering parameter information available, participants had to rely solely on the effective cloud particle size profile. On the other hand, for a number of cases, aerosol radiative effect errors were larger than expected, given the unambiguous and simplistic specification of the scattering parameters.

[40] While performance in the LW CO_2 forcing calculations was good overall with one or two exceptions, estimation of SW CO_2 forcing was found to be a capability that some codes either did not have at all or was restricted to downwelling SFC fluxes. For codes that had full SW CO_2

forcing capabilities, performance was often quite poor, echoing previous findings by *Collins et al.* [2006].

[41] We did not explore in this study whether the spread of the participating models from the reference LBL calculations correlates (inversely) with their computational efficiency. A proper assessment of computational efficiency would have required a common computational platform and therefore significant additional effort. It may be addressed in future phases of CIRC.

[42] A legitimate issue for debate is whether CIRC can assess the degree of progress by approximate RT codes since the seminal ICRCCM study. Alas, it is in practice difficult to quantify precisely the progress or lack thereof for several reasons, such as ICRCCM's much greater number of participating models, ICRCCM cases that were potentially sensitive to a different set of possible model deficiencies, and the fact that ICRCCM included experiments of pure atmospheres with a single absorber. Nonetheless, it appears that the current generation of RT codes does indeed perform better than the codes of the ICRCCM era. If one considers, for example, the LW downwelling flux at the SFC, from the 77 available error values for our 7 clear-sky cases (including subcases) and 11 participating models, only on two occasions (incidentally from the same model) did the error exceed 2%; 53 times the error is within ±0.5%. Contrast this with the results shown in Figure 4 of *Ellingson et al.* [1991] for a midlatitude summer atmosphere, where 12 out of 39 models had errors outside ±2%. While in ICRCCM 20 out of 36 models had flux divergence errors above 2% for the same midlatitude summer atmosphere (9 above 6%), for CIRC only 10 out of 77 times did the error exceed 2%; 51 out 77 times the error remained within ±1%. The total SW downwelling flux at the SFC in CIRC is also simulated quite well by the participating models for the 7 subcases of pristine atmospheres of purely gaseous composition: with one exception where the error is slightly above 7%, none of the 13 models ever produced errors greater than 3%. As already mentioned, the column absorption was not as good, even for the pristine atmospheres: one model (Model 3) produced errors in excess of 10% for all cases, and there was also a case where Model 11 produced a 19% error; errors ~5% were not uncommon, but beyond the above exceptions, SW absorption errors remained below 10%. The ICRCCM figures reported by *Fouquart et al.* [1991] for different standard atmospheres without clouds and aerosols indicate notably worse performance. For downwelling total flux at the SFC the range of errors spanned 5% to 18%, while for column absorbed flux the range of errors had a span that stretched from 16% to 42%, depending on atmosphere and solar geometry.

[43] CIRC's goal, to document the performance of RT models relative to LBL standards, is rather clear-cut. The effort needs to be expanded to include judicious evaluation of model performance in terms of the accuracy needed to address operational GCM requirements for current and future climate simulations. For codes used to generate flux products based on observed input, different performance standards (e.g., minimum accuracy requirements) may be needed depending on specific applications, such as model accuracy being better than the flux range corresponding to the uncertainty of the inputs. Although it may not be an easy task, a key next step would be to attempt to reach a

community consensus on acceptable levels of RT model performance via thresholds for performance metrics such as those presented in Figures 5a and 5b. As CIRC moves to subsequent phases and reassesses RT models with a more expansive set of cases, we hope to converge toward a determination of such application-dependent performance targets.

Appendix A

A1. Input

[44] For the clear-sky CIRC Phase I cases, the atmospheric column is discretized in layers of varying physical thickness, ranging from 54 m near the surface to 4 km for the uppermost layers. For the cloudy cases, the discretization of the cloudy layers is 62.5 m between 375 m and 1 km and 100 m above 1 km. The above layering structure comes from BBHRP where profiles of temperature and humidity are based on radiosonde information up to ~20 km and on climatology at higher levels. The exception is Case 3 for which the profiles throughout the atmospheric column come from ARM's Merged Sounding Value Added Product [*Troyan*, 2010]. The surface (skin) temperature is estimated by inverting the observed broadband LW upwelling surface irradiance using the Stefan-Boltzmann equation assuming unity emissivity throughout the LW spectral range. The air temperature of the lowest atmospheric level is derived from AERI measurements in the opaque 675–680 cm^{-1} region. Each layer's water vapor amount from the sonde measurement is scaled by the ratio of the total precipitable water vapor retrieved from MicroWave Radiometer (MWR) measurements and the sonde's total precipitable water vapor. For the NSA cases, because of the high uncertainty of the MWR measurements in dry conditions, the sonde precipitable water is instead scaled to agree with the value retrieved from NSA extended-range AERI measurements between 535 and 560 cm^{-1}. A CO_2 mixing ratio of 360 ppmv is used for the year 2000 calculations (SGP cases; Cases 1–3, and 6), 375 ppmv is used for the 2004 NSA nominal CO_2 case (Case 4), 750 ppmv is used for the NSA doubled CO_2 case (Case 5), and 380 ppmv is used for the 2005 PYE case (Case 7). The ozone profiles come from monthly means of the Model of Ozone And Related Tracers chemical transport model climatology, after the model's total ozone column amount was scaled to agree with that measured by the Total Ozone Mapping Spectrometer instrument. The AERI-LBL comparison for Case 1 improved in the 9.6 μm band when October 2000 ozone was used, we thus adopt October instead of September ozone values for this case. For all other species, mixing ratios are taken from the U.S. Standard Atmosphere.

[45] The cloud properties of Case 6 are based on ARM's Active Remotely Sensed Clouds Locations (ARSCL) product [*Clothiaux et al.*, 2000], which provides height distributions of hydrometeor reflectivity (and cloud boundaries) every 10 s based on observations from a Millimeter Cloud Radar and Micropulse Lidar. These ARSCL products are combined with thermodynamic profiles from radiosondes and column integrated water vapor estimates from the MWR and inserted into the Microbase cloud property retrieval algorithm [*Miller et al.*, 2003; *U.S. Department of Energy*,

2006], which computes a time-height grid of the liquid water concentration, liquid effective radius, ice water concentration, and ice effective radius (neither of the latter two is pertinent here). In the Microbase retrieval, the initial liquid water content (LWC) profile is integrated to produce an estimate of the Liquid Water Path (LWP); the final LWC profile is obtained by scaling all values by the ratio of the LWP retrieved from coincident MWR measurements and the Microbase LWP estimate. The retrieved cloud properties for each time and height are averaged over a 20 min interval empirically determined to, on average, correspond best with the portion of the cloud fields affecting the 5 min average of irradiance measurements used for radiative closure comparisons. For Case 7 (Pt. Reyes), the cloud property retrievals are based on the MIXCRA inversion algorithm [*Turner*, 2005, 2007]. The cloud is assumed to be vertically homogeneous and its top and base are determined from measurements of a W band Doppler radar operating at 95 GHz. The cloud layer extends from 500 to 2100 above sea level for Case 6 and from 150 m to 270 m for Case 7.

[46] For the SGP clear-sky cases (Cases 1–3) the aerosol optical depths are derived from spectral solar irradiance measurements of the MultiFilter Rotating Shadowband Radiometer (MFRSR) at 6 wavelengths <1 μm. These MFRSR measurements, averaged over 5 min, are used to derive an Angstrom relationship. The aerosol single-scattering albedo is retrieved from the diffuse-to-direct ratio measured in two visible MFRSR channels. The aerosol asymmetry factor is derived from measurements of the backscattered radiation by the surface-based Aerosol Observation System, adjusted to the relative humidity at the surface. Aerosols are assumed to occupy only the lowest six atmospheric layers for the SGP clear-sky cases. For the NSA cases (Cases 4–5) aerosol properties are obtained from ARM's Aerosol Best Estimate (ABE) product. For all clear-sky cases, the single-scattering albedo and asymmetry factor are vertically and spectrally invariant. ABE is also the source of aerosol information for cloudy Case 6, and, while spectral invariance of optical properties is also imposed as in the clear cases, aerosols occupy additional layers above the first 6 layers and their extinction is allowed to vary vertically. With no aerosol measurements of any kind available, the aerosol loading is set to zero for Case 7.

[47] The SGP surface albedo values used in the solar calculations are based on measurements from two downward looking MFRs (essentially the head of the MFRSR instrument), deployed on two towers (at heights of 10 m and 25 m, respectively) at the SGP site. The ratios of the best estimate of upward irradiance from the site's two MFRs to the downward irradiance from the MFRSR is computed for each of the 6 channels of the instrument; these ratios are used to classify the surface type under the towers. This classification, the six ratio values, and published spectral albedos for the particular surface types are then used to create a (piecewise continuous) spectral albedo function. The surface albedo functions associated with each tower are averaged with equal weights to obtain the spectral surface albedo that is input to the reference high-resolution calculation (see *McFarlane et al.* [2011] for further details). For NSA only a single tower is available for surface classification and albedo estimation. Based on a satellite (Multispectral Thermal Imager) image analysis of the region surrounding the NSA

site on the day corresponding to Case 4, it was determined that an appropriate surface albedo would require taking the weighted average of the surface albedo below the tower (with a 85% weight) and the albedo of open water (with a 15% weight). For Case 7 (PYE), MODIS-derived surface (i.e., land) albedos are converted to pseudo-MFR albedos that are used to obtain a spectral albedo function in the exact same manner as for the SGP cases. For all cases, the surface reflectance is assumed to be Lambertian. Since different radiation codes have their own band structure, spectral surface albedos are provided at relatively high spectral resolution (1 cm^{-1}). Two additional spectral functions were provided, the product of surface albedo and the extraterrestrial spectral solar irradiance [*Kurucz*, 1994] in the 0.2–12.2 μm range, and the product of surface albedo and the 1 cm^{-1} downwelling surface irradiances from our reference LBL SW calculations. These spectral functions allow participants to use either unweighted or weighted (by either the TOA or SFC spectral flux) albedos within the (wide) spectral bands appropriate for their RT code.

[48] Finally, the participants are provided with the downwelling broadband flux at the TOA calculated from the [*Kurucz*, 1994] spectral solar irradiance for the Sun-Earth distance and solar zenith angle at the day and time of each case, consistent with the downwelling TOA broadband of the SW LBL calculations. Using these TOA flux values ensures that any errors in the SW flux calculations are only due to differences in the calculated reflective and transmissive properties of the surface-atmosphere system.

A2. Reference LBL Calculations

[49] The reference LW radiative transfer calculations are performed with the Line-by-Line Radiative Transfer Model (LBLRTM) v. 11.1 [*Clough et al.*, 2005]. For the SW, CHARTS v. 4.04 [*Clough et al.*, 2005] is used with gaseous optical depths computed by LBLRTM. The spectroscopic inputs to the model come from the AER_v_2.0 line parameter file and the MT_CKD_2.0 water vapor continuum model. LBLRTM is run using 6 streams (i.e., 3 angles for each of the upward and downward direction) while CHARTS is run using 16 streams and delta M scaling. CHARTS in its current configuration can provide fluxes at only one level per run (we chose TOA and SFC for our purposes). The Planck function within a vertically inhomogenous layer is treated as varying linearly with layer optical depth. The spectral range of the LW calculations (comprising only Earth sources) is 10 to 3250 cm^{-1} (3.077 to 1000 μm) while that for the SW calculations (comprising only solar sources) is 820 to 50000 cm^{-1} (0.2 to 12.195 μm). The cloud optical properties (extinction coefficient, single-scattering albedo and Mie phase function) for the two cloudy cases are calculated from the retrieved layer water contents and effective radii using Mie theory [*Wiscombe*, 1980] applied on an assumed gamma droplet size distribution with a spectral dispersion of 0.12. For LW calculations cloud scattering is not accounted for, so only the absorption part of the cloud extinction optical thickness is retained. Aerosols are treated as in BBHRP, i.e., the spectral aerosol optical thickness is obtained from the Angstrom relationship while the asymmetry factor and single-scattering albedo remain spectrally invariant. The aerosol information used in the CHARTS

calculation is therefore exactly the same as that provided as input to the participants. The aerosol phase function is assumed to follow the Henyey-Greenstein function. Aerosol effects are ignored in the LW. The surface albedo is resolved at 1 cm^{-1} (as that provided to the participants) and is linearly interpolated to the wave number of the calculation. Finally, the surface emissivity is set to unity across the LW spectrum. The output fluxes of the calculations are integrated using a boxcar function into 1 cm^{-1} wide bins.

A3. Radiative Observations

[50] The LW and SW surface observed irradiances for the CIRC SGP and NSA cases are 5 min averages of values provided by ARM Best Estimate Radiative Flux product as included in the BBHRP v_1.4.1 and BBHRP v_1.4.1tK (Case 3) data sets. The 5 min averaging window is centered at the time provided for each case. For the PYE case (Case 7) the LW irradiance measurement value is the mean of the 5 min averages from two pyrgeometers at the site. The PYE SW value is taken from the 5 min average of measurements from the single collocated shaded pyranometer.

[51] The observed TOA irradiance values for Case 4 are from the spatially and temporally closest measurement of the Clouds and the Earth's Radiant Energy System (CERES) measurement. For Cases 3 and 7, the broadband TOA fluxes are inferred from the spatially and temporally closest Geostationary Operational Environmental Satellite (GOES) measurement using a radiance-to-flux and narrowband-to-broadband conversion algorithms [*Minnis and Smith*, 1998]. For the remaining SGP cases the weighted average of the two spatially closest GOES-inferred fluxes temporally bracketing the nominal BBHRP radiative calculation are used. The uncertainty of the GOES-derived TOA fluxes is ~3% for the LW and ~7% for the SW [*Khaiyer et al.*, 2006].

[52] As mentioned in section 2.1, for all cases except Case 5, spectral comparisons between LBLRTM calculations and AERI radiance observations, which are coincident and collocated with the broadband irradiance measurements used in CIRC, establish that the agreement between the surface longwave broadband measurements and reference calculations is supported by available spectral observations. Measurements from the AERI instrument used for the SGP cases (Cases 1–3, 6) and PYE case (Case 7) are produced every ~8 min; all measurements occurring within a 20 min window centered at the time of the broadband measurement are averaged in this analysis. The measurements used for the NSA case (Case 4) are from a "rapid sample" AERI, which provides observations every 20 s that are then averaged over a 20 min window. For the SGP and NSA AERI measurements, a small adjustment is made to the AERI measurements [*Delamere et al.*, 2010] to account for a persistent bias between calculations and AERI measurements at very low radiance conditions.

their respective leaders, C. Kummerow (GRP) and R. Cahalan (IRC), for their guidance. We would like to acknowledge the large number of people involved in CIRC indirectly by developing the ARM data products used and, specifically, M. Miller for leading the cloud retrieval effort used in BBHRP, D. Turner and C. Chiu for CIRC Case 7 cloud and surface input, B. Zak for providing satellite images relevant to CIRC Cases 4 and 5, and M. Khaiyer and P. Minnis for the GOES and CERES satellite-based radiative fluxes.

References

Ackerman, T. P., D. M. Flynn, and R. T. Marchand (2003), Quantifying the magnitude of anomalous solar absorption, *J. Geophys. Res.*, *108*(D9), 4273, doi:10.1029/2002JD002674.

Briegleb, B. P. (1992), Delta-Eddington approximation for solar radiation in the NCAR Community Climate Model, *J. Geophys. Res.*, *97*, 7603–7612, doi:10.1029/92JD00291.

Cagnazzo, C., E. Manzini, M. A. Giorgetta, P. M. De, F. Forster, and J. J. Morcrette (2007), Impact of an improved shortwave radiation scheme in the MAECHAM5 General Circulation Model, *Atmos. Chem. Phys.*, *7*, 2503–2515, doi:10.5194/acp-7-2503-2007.

Chou, M.-D., and M. J. Suarez (1999), A solar radiation parameterization for atmospheric studies, *NASA Tech. Memo.*, TM-1999-10460, vol. 15, 42 pp.

Chou, M.-D., M. J. Suarez, C.-H. Ho, M. M.-H. Yan, and K. T. Lee (1998), Parameterizations for cloud overlapping and shortwave single-scattering properties for use in general circulation and cloud ensemble models, *J. Clim.*, *11*, 202–214, doi:10.1175/1520-0442(1998)011<0202:PFCOAS>2.0.CO;2.

Chou, M.-D., M. J. Suarez, X.-Z. Liang, and M. M.-H. Yan (2001), A thermal infrared radiation parameterization for atmospheric studies, *NASA Tech. Memo.*, TM-2001-104606, vol. 19, 55 pp.

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner (2000), Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites, *J. Appl. Meteorol.*, *39*, 645–665, doi:10.1175/1520-0450(2000)039<0645:ODOCHA>2.0.CO;2.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown (2005), Atmospheric radiative transfer modeling: A summary of the AER codes, *J. Quant. Spectrosc. Radiat. Transfer*, *91*, 233–244, doi:10.1016/j.jqsrt.2004.05.058.

Collins, W. D. (2001), Parameterization of generalized cloud overlap for radiative calculations in general circulation models, *J. Atmos. Sci.*, *58*, 3224–3242, doi:10.1175/1520-0469(2001)058<3224:POGCOF>2.0.CO;2.

Collins, W. D., et al. (2004), Description of the NCAR Community Atmosphere Model (CAM 3.0), *NCAR Tech. Note NCAR/TN-464+STR*, 226 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.

Collins W. D., et al. (2006), Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4), *J. Geophys. Res.*, *111*, D14317, doi:10.1029/2005JD006713.

Collins, W. D., J. K. Hackney, and D. P. Edwards (2002), An updated parameterization for infrared emission and absorption by water vapor in the National Center for Atmospheric Research Community Atmosphere Model, *J. Geophys. Res.*, *107*(D22), 4664, doi:10.1029/2001JD001365.

Costa, S. M. S., and K. P. Shine (2006), An estimate of the global impact of multiple scattering by clouds on outgoing long-wave radiation, *Q. J. R. Meteorol. Soc.*, *132*, 885–895, doi:10.1256/qj.05.169.

Delamere, J. S., S. A. Clough, V. H. Payne, E. J. Mlawer, D. D. Turner, and R. R. Gamache (2010), A far-infrared radiative closure study in the Arctic: Application to water vapor, *J. Geophys. Res.*, *115*, D17106, doi:10.1029/2009JD012968.

Edwards, J. (1996), Efficient calculation of infrared fluxes and cooling rates using the two-stream equations, *J. Atmos. Sci.*, *53*, 1921–1932, doi:10.1175/1520-0469(1996)053<1921:ECOIFA>2.0.CO;2.

Edwards, J. M., and A. Slingo (1996), Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, *Q. J. R. Meteorol. Soc.*, *122*, 689–719, doi:10.1002/qj.49712253107.

Ellingson, R. G., and Y. Fouquart (1991), The intercomparison of radiation codes in climate models: An overview, *J. Geophys. Res.*, *96*, 8925–8927, doi:10.1029/90JD01618.

Ellingson, R. G., and W. J. Wiscombe (1996), The Spectral Radiance Experiment (SPECTRE): Project description and sample results, *Bull. Am. Meteorol. Soc.*, *77*, 1967–1985, doi:10.1175/1520-0477(1996)077<1967:TSREPD>2.0.CO;2.

Ellingson, R. G., J. Ellis, and S. Fels (1991), The intercomparison of radiation codes used in climate models: Long wave results, *J. Geophys. Res.*, *96*, 8929–8953, doi:10.1029/90JD01450.

Fomin, B. A. (2004), A k-distribution technique for radiative transfer simulation in inhomogeneous atmosphere: 1. FKDM, fast k-distribution model for the longwave, *J. Geophys. Res.*, *109*, D02110, doi:10.1029/2003JD003802.

Fomin, B. A. (2006), Monte-Carlo algorithm for line-by-line calculations of thermal radiation in multiple scattering layered atmospheres, *J. Quant. Spectrosc. Radiat. Transfer*, *98*, 107–115, doi:10.1016/j.jqsrt.2005.05.078.

Fomin, B. A., and M. P. Correa (2005), A k-distribution technique for radiative transfer simulation in inhomogeneous atmosphere: 2. FKDM, fast k-distribution model for the shortwave, *J. Geophys. Res.*, *110*, D02106, doi:10.1029/2004JD005163.

Fomin, B. A., and I. P. Mazin (1998), Model for an investigation of radiative transfer in cloudy atmosphere, *Atmos. Res.*, *47–48*, 127–153, doi:10.1016/S0169-8095(98)00056-8.

Fouquart, Y., and B. Bonnel (1980), Computations of solar heating of the Earth's atmosphere: A new parameterization, *Beitr. Phys. Atmos.*, *53*, 35–62.

Fouquart, Y., B. Bonnel, and V. Ramaswamy (1991), Intercomparing shortwave radiation codes for climate studies, *J. Geophys. Res.*, *96*, 8955–8968, doi:10.1029/90JD00290.

Fu, Q., and K.-N. Liou (1992), On the correlated k-distribution method for radiative transfer in nonhomogenous atmospheres, *J. Atmos. Sci.*, *49*, 2139–2156, doi:10.1175/1520-0469(1992)049<2139:OTCDMF>2.0.CO;2.

Fu, Q., K. Liou, M. Cribb, T. Charlock, and A. Grossman (1997), On multiple scattering in thermal infrared radiative transfer, *J. Atmos. Sci.*, *54*, 2799–2812, doi:10.1175/1520-0469(1997)054<2799:MSPITI>2.0.CO;2.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. Collins (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res.*, *113*, D13103, doi:10.1029/2008JD009944.

Illingworth, A. J., et al. (2007), Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, *Bull. Am. Meteorol. Soc.*, *88*, 883–898, doi:10.1175/BAMS-88-6-883.

Jin, Z., T. P. Charlock, K. Rutledge, K. Stamnes, and Y. Wang (2006), An analytical solution of radiative transfer in the coupled atmosphere–ocean system with rough surface, *Appl. Opt.*, *45*, 7443–7455, doi:10.1364/AO.45.007443.

Khaiyer, M. M., D. R. Doelling, P. K. Chan, M. L. Nordeen, R. Palikonda, Y. Yi, D. N. Phan, and P. Minnis (2006), Derivation of improved surface and top of the atmosphere broadband shortwave and longwave fluxes over Atmospheric Radiation Measurement program domains, paper presented at Sixteenth ARM Science Team Meeting, Atmos. Radiat. Meas. Program, Albuquerque, N. M. [Available at http://www.arm.gov/publications/proceedings/conf16/extended_abs/khaiyer_mm.pdf.]

Knuteson, R. O., et al. (2004), The Atmospheric Emitted Radiance Interferometer. Part I: Instrument design, *J. Atmos. Oceanic Technol.*, *21*, 1763–1776, doi:10.1175/JTECH-1662.1.

Kurucz, R. L. (1994), Synthetic infrared spectra, in *Infrared Solar Physics: Proceedings of the 154th Symposium of the International Astronomical Union*, vol. 154, edited by D. M. Rabin, J. T. Jefferies, and C. Lindsey, pp. 523–531, Kluwer Acad., Dordrecht, Netherlands.

Li, J. (2002), Accounting for unresolved clouds in a 1D infrared radiative transfer model. Part I: Solution for radiative transfer, including cloud scattering and overlap, *J. Atmos. Sci.*, *59*, 3302–3320, doi:10.1175/1520-0469(2002)059<3302:AFUCIA>2.0.CO;2.

Li, J., and H. W. Barker (2002), Accounting for unresolved clouds in a 1D infrared radiative transfer model. Part II: Horizontal variability of cloud water path, *J. Atmos. Sci.*, *59*, 3321–3339, doi:10.1175/1520-0469(2002)059<3321:AFUCIA>2.0.CO;2.

Li, J., and H. W. Barker (2005), A radiation algorithm with correlated-k distribution. Part I: Local thermal equilibrium, *J. Atmos. Sci.*, *62*, 286–309, doi:10.1175/JAS-3396.1.

Li, J., S. Dobbie, P. Räisänen, and Q. Min (2005), Accounting for unresolved clouds in a 1-D solar radiative-transfer model, *Q. J. R. Meteorol. Soc.*, *131*, 1607–1629, doi:10.1256/qj.04.31.

McFarlane, S. A., K. L. Gaustad, E. J. Mlawer, C. N. Long, and J. Delamere (2011), Development of a high spectral resolution surface albedo product for the ARM Southern Great Plains central facility, *Atmos. Meas. Tech.*, *4*, 1713–1733, doi:10.5194/amt-4-1713-2011.

Miller, M. A., K. L. Johnson, D. T. Troyan, E. E. Clothiaux, E. J. Mlawer, and G. G. Mace (2003), ARM value-added cloud products: Description and status, paper presented at Thirteenth ARM Science Meeting, Atmos. Radiat. Meas. Program, Broomfield, Colo. [Available at http://www.arm.gov/publications/proceedings/conf13/extended_abs/miller-ma.pdf.]

Minnis, P., and W. L. Smith Jr. (1998), Cloud and radiative fields derived from GOES-8 during SUCCESS and the ARM-UAV spring 1996 flight series, *Geophys. Res. Lett.*, *25*, 1113–1116, doi:10.1029/98GL00301.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *J. Geophys. Res.*, *102*, 16,663–16,682, doi:10.1029/97JD00237.

Mlawer, E. J., P. D. Brown, S. A. Clough, L. C. Harrison, J. J. Michalsky, P. W. Kiedron, and T. Shippert (2000), Comparison of spectral direct and diffuse solar irradiance measurements and calculations for cloud-free conditions, *Geophys. Res. Lett.*, *27*, 2653–2656, doi:10.1029/2000GL011498.

Mlawer, E. J., et al. (2002), The Broadband Heating Rate Profile (BBHRP) VAP, paper presented at Twelfth ARM Science Meeting, Atmos. Radiat. Meas. Program, St. Petersburg, Fla. [Available at http://www.arm.gov/publications/proceedings/conf12/extended_abs/mlawer-ej.pdf.]

Moncet, J.-L., and S. A. Clough (1997), Accelerated monochromatic radiative transfer for scattering atmospheres: Application of a new model to spectral radiance observations, *J. Geophys. Res.*, *102*, 21,853–21,866, doi:10.1029/97JD01551.

Oreopoulos, L., and E. Mlawer (2010), The Continual Intercomparison of Radiation Codes (CIRC): Assessing anew the quality of GCM radiation algorithms, *Bull. Am. Meteorol. Soc.*, *91*, 305–310, doi:10.1175/2009BAMS2732.1.

Räisänen, P., and H. W. Barker (2004), Evaluation and optimization of sampling errors for the Monte Carlo Independent Column Approximation, *Q. J. R. Meteorol. Soc.*, *130*, 2069–2085, doi:10.1256/qj.03.215.

Shephard, M. W., S. A. Clough, V. H. Payne, W. L. Smith, S. Kireev, and K. E. Cady-Pereira (2009), Performance of the line-by-line radiative transfer model (LBLRTM) for temperature and species retrievals: IASI case studies from JAIVEx, *Atmos. Chem. Phys.*, *9*, 7397–7417, doi:10.5194/acp-9-7397-2009.

Stokes, G. M., and S. E. Schwartz (1994), The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed, *Bull. Am. Meteorol. Soc.*, *75*, 1201–1221, doi:10.1175/1520-0477(1994)075<1201:TARMPP>2.0.CO;2.

Tarasova, T. A., and B. A. Fomin (2007), The use of new parameterizations for gaseous absorption in the CLIRAD-SW solar radiation code for models, *J. Atmos. Oceanic Technol.*, *24*, 1157–1162, doi:10.1175/JTECH2023.1.

Troyan, D. (2010), Merged sounding value-added product, *Tech. Rep. DOE/SC-ARM/TR-087*, 13 pp., U.S. Dep. of Energy, Washington, D. C. [Available at http://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-087.pdf.]

Turner, D. D. (2005), Arctic mixed-phase cloud properties from AERI lidar observations: Algorithm and results from SHEBA, *J. Appl. Meteorol.*, *44*, 427–444, doi:10.1175/JAM2208.1.

Turner, D. D. (2007), Improved ground-based liquid water path retrievals using a combined infrared and microwave approach, *J. Geophys. Res.*, *112*, D15204, doi:10.1029/2007JD008530.

Turner, D. D., et al. (2004), The QME AERI LBLRTM: A closure experiment for downwelling high spectral resolution infrared radiance, *J. Atmos. Sci.*, *61*, 2657–2675, doi:10.1175/JAS3300.1.

U.S. Department of Energy (2006), Continuous profiles of cloud microphysical properties for the fixed atmospheric radiation measurement sites, *Rep. DOE/SC-ARM/P-0609*, 7 pp., Washington, D. C.

Wiscombe, W. J. (1980), Improved Mie scattering algorithms, *Appl. Opt.*, *19*, 1505–1509, doi:10.1364/AO.19.001505.

Zhang, Y., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko (2004), Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data, *J. Geophys. Res.*, *109*, D19105, doi:10.1029/2003JD004457.

J. Cole and J. Li, Canadian Centre for Modeling and Analysis, Environment Canada, 4905 Dufferin St., Downsview, ON M3H5T4, Canada.

J. Delamere, Tech-X Corporation, 5621 Arapahoe Ave., Boulder, CO 80303, USA.

B. Fomin, Central Aerological Observatory, 3 Pervomayskaya St., Dolgoprudny, 141700 Russia.

M. Iacono and E. Mlawer, Atmospheric and Environmental Research Inc., 131 Hartwell Ave., Lexington, MA 02421-3126, USA.

Z. Jin and F. Rose, Science Systems and Applications Inc., 1 Enterprise Pkwy., Ste. 200, Hampton, VA 23666, USA.

J. Manners, Met Office, FitzRoy Road, Exeter, UK.

L. Oreopoulos, Climate and Radiation Laboratory, NASA Goddard Space Flight Center, Mail Code 613, Greenbelt, MD 20771, USA. (lazaros.oreopoulos@nasa.gov)

P. Räisänen, Finnish Meteorological Institute, PO Box 503, FI-00101 Helsinki, Finland.

W. B. Rossow, Cooperative Remote Sensing Science and Technology Center, City College of New York, 2880 Broadway, New York, NY 10025, USA.

T. Shippert, Pacific Northwest National Laboratory, PO Box 999/MS K9-24, Richland, WA 9935, USA.

M. J. Wilson, I.M. Systems Group, Ste. 204, Airmen Memorial Building, 5211 Auth Rd., Camp Springs, MD 20746, USA.

Y. Zhang, RM 320-B, Department of Applied Physics and Applied Mathematics, Columbia University, 2880 Broadway, New York, NY 10025, USA.